

Open Research Online

The Open University's repository of research publications and other research outputs

'Difficulty' In Text As A Function Of Syntactic Complexity: A Study Of Syntactic Complexity Within And Between Sentences

Thesis

How to cite:

Edwards, Neville (1980). 'Difficulty' In Text As A Function Of Syntactic Complexity: A Study Of Syntactic Complexity Within And Between Sentences. MPhil thesis The Open University.

For guidance on citations see [FAQs](#).

© 1980 Neville Edwards



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000f7e2>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

'DIFFICULTY' IN TEXT AS A FUNCTION OF SYNTACTIC COMPLEXITY:

A study of syntactic complexity within and between sentences.

Submitted as a requirement for the degree of Master of
Philosophy to the Open University.

March 1980

Neville Edwards

The Textual Communication Research Group

Institute of Educational Technology

The Open University

Walton Hall

Milton Keynes

Date of submission: 26.3.80

Date of award: 1.7.80

ProQuest Number: 27777431

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 27777431

Published by ProQuest LLC (2020). Copyright of the Dissertation is held by the Author.

All Rights Reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

A B S T R A C T

This work proposes possible categories for analysing complexity in English text. Particular attention is paid to the following categories: the long subject construction, the nominal group, the appositional group, and interpolation structures. These categories are analysed in relation to two dimensions of text: firstly, in relation to the individual sentence, and secondly, in relation to the larger discourse context of clause relations. (A clause relation is a binary relation between members. Each member consists of one or more clauses or sentences).

The first dimension examines how the presence of these categories functions to interrupt the structure of the clause as a minimal grammatical form. It is assumed that the structure of the clause is basic to the structure of text and is central to the reading process. When the structure of the clause is interrupted excessively by these categories, it is predicted that there will be a consequent impairment of reading comprehension.

The second dimension of complexity examines how the structural interruption of the individual sentence can entail the structural interruption of the clause relation. Syntactic complexity at this level of structure is increased when the minimal grammatical form of the clause relation (that is, the relation between two members) is delayed excessively by structural interruption.

Thus, although the individual clause is the primary unit of description in this analysis of complexity, the analysis is not restricted to the single sentence. Moreover, the individual categories proposed are descriptive procedures which derive from the examination of texts as aspects of performance, not as aspects of competence. Finally, the notion of clause is put forward as a grammatical constraint which has important implications for our production of texts. The clause is fundamental to the structure of texts and provides the mechanism for its cohesion.

CONTENTS

Abstract	i
Foreword	iii
Acknowledgements	vi
<u>INTRODUCTION</u>	
General background	1
Scope and Approach of the Present Work	21
<u>CHAPTER ONE: The Long Subject Construction</u>	
Section 1.1	29
Section 1.2	36
Section 1.3	46
<u>CHAPTER TWO:</u>	
Section 2.1 The Structure of Nominal Groups	51
Section 2.2 Apposition	61
<u>CHAPTER THREE: Interpolation</u>	
Section 3.1 General Criteria	84
Section 3.2 Examples of Interpolation	93
<u>CHAPTER FOUR: Clause Relations</u>	
	114
<u>CHAPTER FIVE: Conclusion</u>	
	135
NOTES	140
REFERENCES	152
PRINCIPAL REFERENCES	157
GLOSSARY OF SOME OF THE TERMS USED	158

FOREWORD

The question of individual text 'difficulty' is of great importance in an educational context where texts are used for instructional purposes. As the term is used here, difficulty is an inherent property of text which varies according to certain aspects of textual organisation. The difficulty levels of text may be grouped, broadly, under three headings: the typographical, the semantic and the grammatical. The typographical design and layout of a text has obvious consequences for the clarity of the presentation of its subject-matter, while the semantic organisation determines the actual form of the conceptual structure. At the same time, the grammatical organisation of a text also determines in part the ease with which the reader can assimilate this conceptual structure. Clearly, all three components overlap and each contributes to the level of difficulty in the text. The research presented in this thesis, however, is almost exclusively concerned with the grammatical component of text structure as a partial determinant of its difficulty level. It is concerned to show how syntactic complexity in text can lead to unnecessary difficulty for the reader. The intention of the work undertaken is to produce a set of helpful insights which contribute to the rapidly expanding field of discourse analysis.

The conclusions of this research may be most relevant to a particular style of reading, namely, where the text is read closely and completely at first sight. (Other styles of reading such as scanning, etc., may or may not be similarly affected by grammatical structure.) Successful reading of the type we are concerned with therefore entails a set of predictive strategies on the part of the reader, a set of expectations about the grammatical structure of the text. It is when these predictive strategies are upset by syntactic complexity that reading performance is impaired. Of course, there are many possible ways of analysing such syntactic complexity, depending on the purpose of the analysis. The main purpose of this research, however, is to provide a replicable set of categories which admit of a practical application. The latter is especially important in that it provides a means of testing the adequacy of the categories themselves.

Inevitably, most methods for analysing the sentence are wedded to particular 'theories' of language structure, which go beyond the raw language data. As a result, the means of description is often too abstract for most practical purposes. In the case of transformational-generative grammar, for example, many of the categories used in the analysis of the sentence are not directly represented in its surface syntactic structure at all. The descriptive scope of such categories, moreover, is limited to the single sentence. In contrast to this, the present approach bases the categories of description on the surface syntactic form of the sentence and is not restricted to the single sentence. As an aspect of the relatively recent developments in text grammar, it is also concerned with supra-sentential relations. Thus, although the individual sentence (or clause) is the primary unit of description, it also provides the basis for analysing grammatical (and semantic) relations between sentences. Accordingly, two dimensions of syntactic organisation in text are chosen for study. The first dimension has to do with grammatical structure at the level of the individual clause, while the second is concerned with grammatical structure at the level of inter-sentential relations, as manifested in clause relations.

The approach used for dimensions of complexity derives from the work in English Grammar by Dr. Eugene Winter of The Hatfield Polytechnic. The analysis for complexity in relation to the grammatical structure of the clause is based on Dr Winter's analysis of the structure of clause. The notion of Interpolation, for example, (See Chapter Three below) is examined in great detail in Winter's forthcoming book "Towards a contextual grammar of English: a study of the written paragraph in English." (George Allen and Unwin, in press). Also, some of the examples used in the present work (e.g. those used in the discussion of the long subject construction in Chapter One) are Dr Winter's own unpublished examples. Even more important, the analysis of semantic relations between sentences

(or clauses) of the second dimension is a derivative account of the work done in Winter (ibid.) and Winter's forthcoming long article "Replacement as a fundamental function of the sentence in context" (Forum Linguisticum Vol.111, No. 2 1980). The notion of Replacement used in my work is taken directly from the latter.

Replacement unifies both dimensions of complexity examined here, which analyze the sentence (or clause) in terms of its contextual role with the 'paragraph'. Inevitably, therefore, syntactic complexity at the level of the individual sentence has implications for the degree of complexity of the grammatical structures beyond the sentence. If the sentence has a diversionary complexity, this may impair the comprehension of the clause relation of which it is a member.

It is hoped that the categories proposed in this work are relevant as (partial) criteria for a certain type of difficulty level in text. It should perhaps be emphasised, however, that these categories are put forward as 'possible' causes of syntactic complexity in text; as such, they must be subjected to empirical testing and possible revision in the light of this testing. Most importantly, they are intended as practical explanations of syntactic complexity on the assumption that the obvious place to start looking for causes of difficulty in comprehension of the sentence is in its syntax.

ACKNOWLEDGEMENTS

I joined the Institute of Educational Technology at the Open University in October 1977 as a member of the Textual Communication Research Group under the supervision of Mr. Michael Macdonald-Ross. The work done by the research group has been particularly concerned to identify likely problem-areas in instructional texts and to develop the analytical frameworks necessary to study such problem-areas. A natural extension of this work has been to try to develop a grammatical framework in which to analyse the structure of texts; more specifically, to provide a grammatical framework whose descriptive scope is not limited to the individual sentence but which permits the analysis of larger discourse structures as well. The work presented in this thesis represents an attempt to provide such a descriptive framework.

The research group was fortunate in having an aim shared by Dr. Eugene Winter's research group at The Hatfield Polytechnic. The interests of the two research groups were united by a common purpose. The technical content of the present work is therefore the outcome of the series of seminars and discussions which I have had with Dr. Winter at The Hatfield Polytechnic between October 1977 and October 1979. I have been able to avail myself of the work done by Dr. Winter in the analysis of English texts in his forthcoming publications. These are, firstly, "Replacement as a fundamental function of the sentence in context" (Forum Linguisticum, Vol 111, No.2 1980); and secondly, "Towards a contextual grammar of English: a study of the written paragraph in English" (George Allen & Unwin, in press). The technical content of the present work derives from both of these works in particular. Dr Winter has shown himself to be an admirable teacher of unlimited energy and great understanding and I am grateful for his help. The shape and direction of this research has been continuously monitored by my supervisor, Mr Michael Macdonald-Ross, throughout this period. The development and progress of this work has thus been guided by

the careful efforts of these two people. Having said this, I nevertheless accept full responsibility for any errors or omissions that this work may contain.

Thanks are also due to my research colleague Mr. P.C. Whalley, who has drawn my attention to many important sources of research most relevant to the topic of study chosen for this thesis. I would like to express my gratitude to Mrs Eleanor Smith, who has been most helpful in providing information about the precise source of references which I was unable to complete myself. Thanks are also due to Mrs Julie Painter, who has painstakingly typed this manuscript and shown great patience with the numerous changes she has had to make to the proofs.

Finally, I gratefully acknowledge the benefit I have received from the seminars which I was allowed to give at The Institute on the topic of this research. I found the resulting discussions most fruitful insofar as they enabled me to test out and revise some of the ideas put forward in this work.

INTRODUCTION

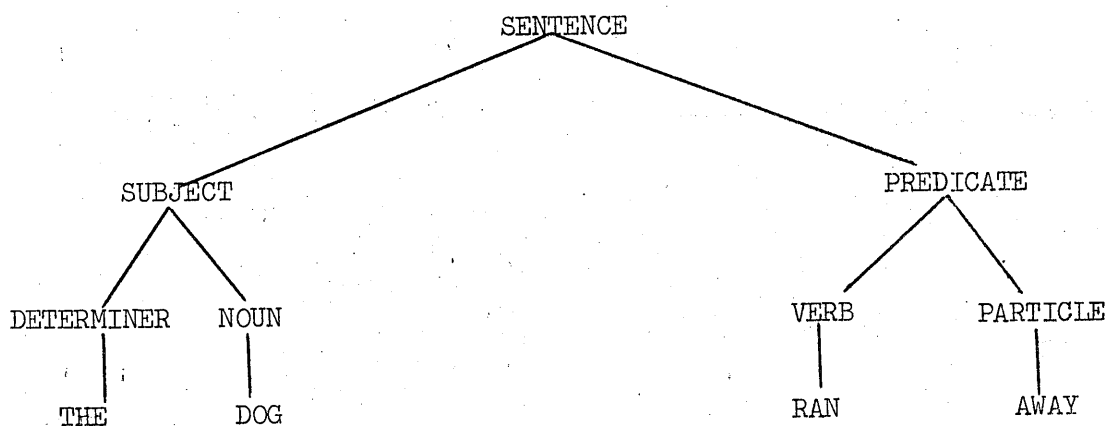
GENERAL BACKGROUND

The aim of the present work is to develop a set of descriptive categories with which to approach the analysis for syntactic complexity in English text. This presupposes that 'difficulty' in text is partly a function of its syntactic organisation. The categories that are put forward are proposed as hypotheses for psychological or statistical validation. The initial formulation of the categories is based on analysis of English text. The notion of 'text' as opposed to 'sentence' is crucial here. 'Text' may be taken to subsume the category of 'sentence' but not conversely. This has important consequences for the scope of the inquiry into difficulty in text. The grammatical analysis employed in the development of these categories rests on an alternative approach to language structure to that proposed in transformational-generative grammar (henceforth T.G.). Unlike the latter, the present approach is part of the current attempt to develop a text grammar of English. Moreover, it takes the text in its real-world context as the primary data for linguistic description and rejects the view that the categories of linguists' grammar need not be subjected to empirical testing. The theory of T.G. nevertheless informs some models of text grammar that have been proposed. In order to show how the present approach (as an aspect of text linguistics), differs from that of T.G., it is necessary to trace the development of syntactic analysis in T.G. from its early stages. T.G. itself, however, must be seen as a reaction against still earlier approaches to linguistic description. In the first part of this introduction, therefore, an attempt will be made to supply the necessary details of this background to recent trends in text linguistics. The main purpose here will be to draw out the implications of alternative theories for difficulty in text and to assess their validity and usefulness.

T.G., as expounded in Noam Chomsky's 'Syntactic Structures (1957) and 'Aspects of the Theory of Syntax' (1965), was largely a reaction against early American structuralism, the main exponent of which was Leonard Bloomfield. In his by now classic 'Language' (1935), Bloomfield attempted to delineate the scope of linguistic description. Rejecting all types of mentalism, he maintained that linguistics is concerned with the 'mechanistic' aspects of language study (physicalism); that is, it is concerned with those aspects of language behaviour which are observable and hence, measureable. Semantics, for example, was defined in behaviourist terms. Verbal behaviour is the result of (Response to)

external influences (or Stimuli). More precisely, language is a substitute - response for physical activity. By the same token, Bloomfield maintained that the meaning of individual lexical items should be defined within terms of the things to which they refer in the real world. Empirical science, he claimed, enables us to define the meaning of lexical items in this way. (Sodium Chloride, for example, is defined as NAC 1). Bloomfield was nevertheless aware of the inherent limitation to this approach. Put simply, this limitation is that a 'complete' semantic description of language in these terms is possible only when we have an accurate scientific knowledge of everything in the speaker's universe. In consequence, Bloomfield held that semantics is a subordinate goal of linguistic inquiry. The primary goal of linguistic description is the taxonomic classification of the syntax of language. The taxonomic range of this classification extends from the phoneme to the complete sentence. However, Bloomfield (1926) regarded the sentence as the largest unit for description in syntactic analysis. The latter he held to be an independent linguistic form which is not included in any larger linguistic form.

Bloomfield's method of approach to syntactic analysis was codified in Immediate Constituent analysis (I.C.). According to this method, a sentence is split into two immediate constituents, which are themselves split up into further constituents. At each cutting point in the sequence an identifying node or label is produced. A sentence is thus seen to be organized into a series of layers of constituents, not merely as a string of elements. For instance, the sentence 'The dog ran away', would be analyzed by I.C. analysis in a form which is made explicit by the following tree:



Bloomfield maintained that the latter was merely a formalization of the method used by native speakers to segment speech. The great advantage of this method was that it was explicit and concise: the sentence, according to this approach, was successively cut into its immediate constituents in a bottom-to-top direction. The transformational approach of 'Syntactic Structures' however, reversed the direction of this segmentation; the sentence here was 'generated' from top to bottom. This latter approach resulted from the failure of I.C. analysis to account for certain observed facts of language. According to the account given in Jacobson (1977), I.C. analysis is an inadequate grammatical framework for three reasons. These are given as follows.

Firstly, I.C. analysis cannot satisfactorily describe discontinuous constituents in the sentence. For example, it cannot describe the following two sentences:

(1) Vi S Vii
Is John coming?

(2) Oind. V A S V
Who was it actually that Jane said

Od.
[you had so generously promised to give the money to?]

(Jacobson p. 13)

In sentence (1) the verb phrase is split, with the operator is occurring in sentence-initial position and the main verb (coming) occurring at the end of the sentence. This sequence is the typical interrogative mood of sentences and illustrates a case where I.C. analysis cannot be applied. Similarly, in sentence (2) the relative pronoun who (which replaces the indirect object) occurs at the beginning of the sentence. The preposition (in this case to) which should co-occur with the indirect object is separated in this sequence and is placed at the end of the sentence. The corresponding declarative form of this sentence is provided by example (3) below:

(3) S V Od. Od. Oind.
Jane said that [you had to give [the money to [X]]]

where 'X' represents the indirect object. As can be seen from example (3), the preposition to immediately precedes the indirect object. The brackets

一
一
一

□ □ □ □

- □ □ □

[
[
[
[

[
[
[
[

- 1

The two possible readings of this sentence are provided in examples (9) and (10) below:

- (9) S V O
 [John's friends] disregarded [him,] which disturbed him.
- (10) S V Vii O
 [John's friends] were disregarding [what disturbed him.]

(Jacobson, p.15)

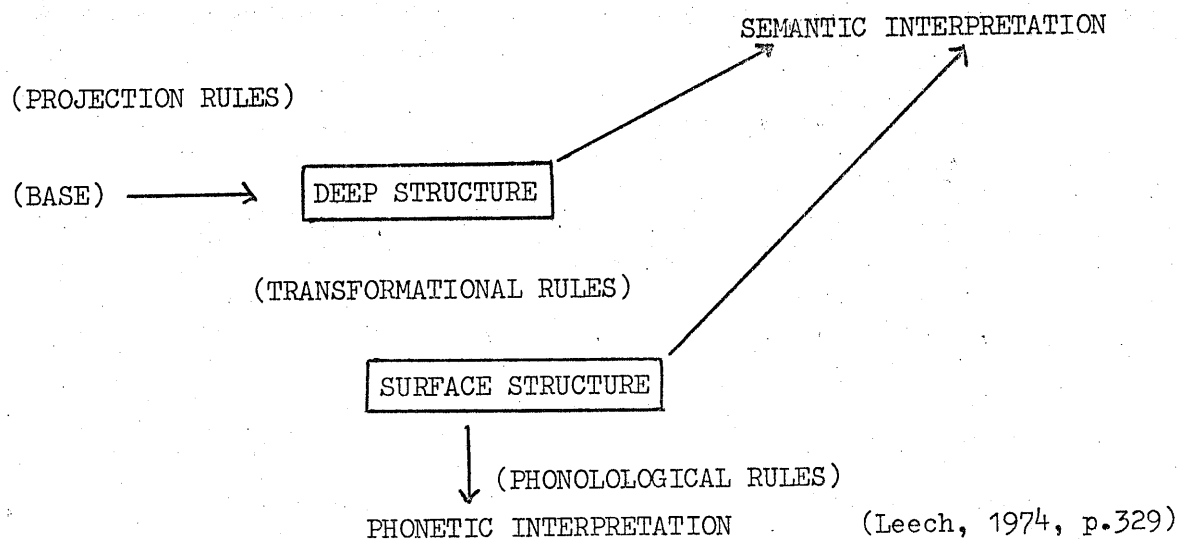
The ambiguity of sentence (8) thus rests on a structural difference of transitivity. In example (9), for example the object of the verb disregarded is John himself (him), while in sentence (10) it is 'what disturbed him'. While this transitivity difference would receive two different parsings in a transformational analysis, I.C. analysis would assign sentence (8) only one structure.¹

Apart from I.C. analysis, Chomsky also rejected Bloomfield's behaviourist position, especially in relation to semantic description.² Against Bloomfield's physicalism, Chomsky based linguistic theory on mentalist doctrine (1965). Chomsky recaptured the Saussurean distinction between la langue and la parole with the corresponding distinction between competence and performance. In striking contrast to Bloomfield's empiricism, Chomsky asserted that the intuitions of native speakers provide the data for linguistic description. In the first version of T.G. (1957), Chomsky was concerned with the grammatical/ungrammatical distinction between sentences. He also attempted to show the syntactic similarities (relatedness) between sentences; here, sentences are transformationally-derived from underlying Kernel sentences. Emphasis was also placed on the recursive aspect of language, where sentences could be of indefinite length and variety. It is interesting that Chomsky nevertheless retained the Bloomfieldian restriction of the sentence as the upper limit to syntactic description. Moreover, in a way reminiscent of Bloomfield, he largely ignored meaning in this early version of the theory.

In the 'classical' theory of Chomsky (1965), the notion of deep structure was introduced.³ According to this modified version of the theory, a sentence is seen to be organized on two syntactic levels: a deep structure and a surface structure. The surface structure of the sentence is derived from the deep structure by means of transformational rules, which involve such operations as deletion of constituents, the insertion of constituents and the movement of constituents to another part of the sentence, etc. The rules specifying the deep structure are rules of the base component (called phrase structure rules).

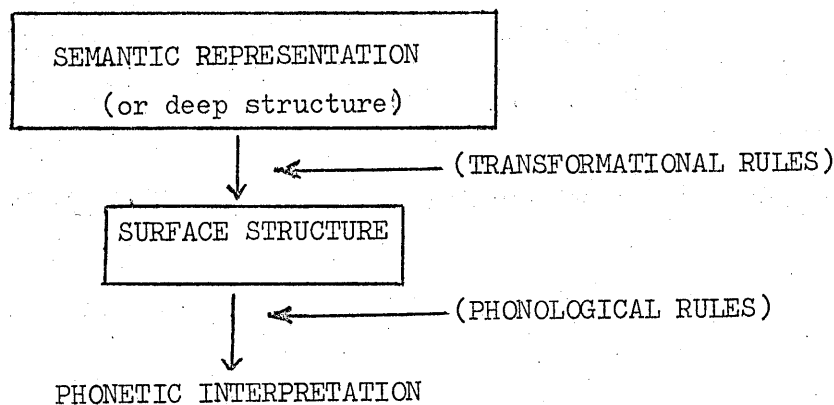
Transformational rules make up the transformational component and produce surface structures. In addition, there are two interpretive components to the grammar: the phonological and the semantic. As a result of the latter, this version of T.G. came to be known as the interpretive semantics position. Here, the phonetic representation of a sentence is derived from its surface structure by means of phonological rules, while the semantic interpretation of a sentence is derived from the deep structure by the operation of projection rules. The important difference between this 1965 version and that of 1957 is the greater importance attached to meaning by the former. This was a consequence of the important article 'The Structure of a Semantic Theory', by Katz and Fodor (1965). In sum, the principal claims of this 'classical' version of T.G. were that syntactic structure is the only level of syntax relevant to the specification of phonetic interpretation; and that syntactic deep structure is the only level of syntax relevant to semantic interpretation. The latter claim assumed that transformations are meaning preserving; that is that they do not alter the meaning of the structures to which they are applied. According to this idea, all sentences with the same deep structure have the same meaning.

Since 1965, however, the 'classical' theory of T.G. has been modified in an important respect. It was noticed that some aspects of meaning (i.e., those involving negation, quantification and information focus) appear to be more the result of the surface structure of a sentence than its deep structure. The theory was therefore changed slightly so that the projection rules which specify meaning should operate on surface structures (and at certain points during a transformational derivation) instead of just on deep structures. In short, this revised interpretive position no longer holds that all sentences with the same deep structure have the same meaning. This new version is illustrated by the following diagram:



Within this new interpretive theory, deep structure is a level justified in large part on syntactic grounds.

Quite a different development of the 1965 version, however, is provided by the Generative Semantics hypothesis. According to this hypothesis, deep structure is held to be so deep that it is 'identical' with the semantic representation of a sentence. Whereas the base component in the interpretive semantics model is syntactic, it is semantic in the generative semantics model. The notion of a sentence's semantic representation is central to theories of semantic description, which attempt to account for the inference-relations between sentences. Consider, for example, the sentence 'John hates to lose'. According to such theories, the semantic representation of this sentence would be: 'John hates [John lose]' (omitting for simplicity the particle 'to'). In this representation, the subject John is the subject of both verbs. The correctness of analyzing the sentence in this way is held to be affirmed by the relations of inference between this sentence and other sentences, where John occurs as the overt subject of the verb lose. This is the case in the inferred sentence 'If John loses, he will be unhappy'. In the generative version of T.G., the deep structure of 'John hates to lose' is also 'John hates [John lose]', on the grounds that John is the subject of the main clause and the subordinate clause before the operation of equi-NP Deletion has applied. In the generative semantics model, therefore, there is no longer the necessity for the projection rules to supply an interpretation of deep structure because the deep structure is the semantic interpretation. This idea is shown more clearly in the following diagram:



(Leech, *ibid*, p. 330).

Although the generative semantics model is simpler in that it eliminates the projection rule component, its disadvantage is that the transformational derivation of a sentence is much longer than in the interpretive semantics model.

These two different models of the representation of a sentence, the generative and interpretive, developed because of the alleged distinction between the base components of a grammar and the derived components which 'interpret' the output of the base component. In the case of interpretive semantics, the semantic representation of a sentence is derived from a syntactic base, while in the generative model the surface syntactic representation of a sentence is derived from a syntactic base. These two different positions have given rise to a vigorous debate in linguistics about the precise form of the grammar. The generative semantics theorists maintain that transformations do not change meaning. This latter claim has attracted severe criticism in view of the way the scope of negation and quantification are conditioned by the ordering of constituents in surface structure. The weak point in the interpretive semantics model, however, is the claim that there is a "valid level of linguistic abstraction corresponding to the deep syntactic structure of classical theory" (Leech, 1974, p.332). Moreover, there is no necessary reason to assume that such deep structure possesses all the properties that the interpretive theory ascribes to it. Syntactic deep structure, for example, is the level where lexical items are inserted, relations of subcategorization are defined, the starting point for transformations and, finally, where the basic constituents of sentences (such as subject and object, etc.) are defined. (Leech, *ibid.*)

Notwithstanding their disagreement on this issue, the generative and interpretive semantics schools are unified by the concept of deep structure and differ only in its interpretation. As we saw earlier, I.C. analysis cannot account for such observed properties of sentences as structural ambiguity, grammatical similarity, discontinuity between constituents, etc. The existence of such observable properties led Chomsky to assume that there must be such a level of deep syntactic structure to explain these properties. The assumption of deep structure on these grounds, however, has been criticized by Botha (1973) in logical terms. Botha suggests that the logic of justification for deep structure may be characterized as regressive reduction. The argument for deep structure, Botha suggests, may be characterized in the following form:

- (B) If a level of deep structure is introduced into the general-linguistic theory, then it follows that sentences can have the grammatical properties $P_1 P_2 \dots, P_n$.

Sentences do have the grammatical properties, $P_1 P_2 \dots, P_n$.

∴ A level of deep structure should be introduced into the general - linguistic theory.

(p. 82)

If we represent the premisses in this argument by propositional letters, then its logical form is made even more clear in the following way:

$$P \supset Q, Q \vdash P$$

This is a logical fallacy known as affirming the consequent. In other words, the consequent (Q) of the first (conditional) premiss is taken as the basis for the inference of the unknown antecedent (P). The deep structure argument, according to Botha, assumes the truth of certain observed grammatical properties of sentences, and infers the notion of deep structure as a possible explanation for the existence of these properties. Furthermore, the content of the conclusion is not associated with the premisses and, therefore, does not follow from the premisses. This means that the argument is an essentially non-demonstrative pattern of inference: it would be quite possible to select some other conclusion from the premisses. For example, from the statement 'If a man is shot directly in the heart, then he will die', we cannot conclude that if some other person dies, then he too must have been shot in the heart; there may be quite a different explanation. By the same token, Botha maintains that the properties explained by deep structure could also be explained by some alternative hypothesis. He concludes that "...transformational grammar will not make significant progress in its search for insight into the nature of natural language(s) by adding to the already large stocks of alternative linguistic hypotheses that fail to be sufficiently justifiable. Real progress towards such insight will not be made unless transformationalists clarify and motivate the diverse justificatory devices which collectively constitute the logic of justification of the field of transformational grammar" (p. 331). This is particularly the case with regard to the generative-interpretive debate.

This contention about the inconclusive nature of the justification for the existence of deep structure would be false if it were possible to psychologically validate the notion. That is, if it were possible to prove its existence by psychological tests. This question of psychological reality applies equally, of course, to the notion of transformations. There are three possible positions to

be taken on the question of the psychological reality of transformations: the isomorphistic, semi-isomorphistic and non-isomorphistic.⁴ According to the isomorphistic argument, the rules (transformations) of linguists' grammars have a direct neurological correlate in the speaker/hearer. In the semi-isomorphistic hypothesis, although the rules have no systematic correspondence in language, such linguistic levels as phonology, syntax etc., do, and speakers proceed from level to level in an ordered fashion. Finally, the non-isomorphistic hypotheses asserts that there is no systematic relationship between the rules of T.G. and language. With respect to the isomorphistic hypothesis, it is not clear whether this would apply equally to speakers and hearers; that is, between encoders and decoders. The position of Chomsky himself on the correspondence hypothesis is not at all clear. However, according to Sampson (1975), Chomsky is claiming that any language which can be defined can be specified by a finite set of formulae of the form $\phi \rightarrow \psi$ (i.e., rules of an unrestricted rewrite system), where ϕ and ψ are sequences of symbols. He is arguing, Sampson maintains, that attested languages fall within a special class of definable languages and the rules which describe them represent an empirical discovery.⁵ Unfortunately, there is as yet no definitive psychological evidence which would enable us to conclusively decide the correspondence hypothesis in favour of any one of the three positions. The view adopted in the present work, however, is that the rules of T.G. and, more specifically, the level of deep structure do not represent psychologically real properties of language; nor are they necessary on logical grounds.

We may, finally, consider the view of Johnson-Laird (1970) on this question of the psychological reality of deep structure. Johnson-Laird claims that although deep structure is supposed to be a precondition for meaning, it nevertheless depends on meaning for its existence in some cases. For instance, he cites the following two sentences:

- (14) John promised the man to escape.
- (15) John persuaded the man to escape.

(p. 267)

He argues that if the listener (or reader) were not familiar with the meaning of promised and persuaded, he would not know that, in the first sentence, the subject of escape is John, but is the man in the second sentence. It will be recalled that Jacobson characterized the failure of

I.G. analysis as its inability to explain differences of this kind; the alleged ability of T.G. to do so was held to be evidence for the existence of deep structure. Johnson-Laird, however, concludes that in view of examples like (14) and (15) deep structure has no independent psychological status and should be seen as "polarized into two separate components" (p.267). The first component is concerned with the parsing of surface structure, while the second is concerned with interpreting lexical words. We may also note that transformational rules have themselves been criticized as non-predictive and un-insightful (Itkonen, 1975), owing to their reliance on native speakers' intuitions as a decision-procedure for the grammaticality of sentences. If Chomskyan grammars cannot, as Itkonen suggests, predict properties of syntax, then it is difficult to see how they can justify the scientific and psychological claims made for them.

It has been necessary to pursue the various ramifications of the generative-interpretive semantics debate and to consider the psychological status of T.G. in order to provide the necessary background to the alleged implications of T.G. for syntactic complexity and comprehension of sentences. In this instance, consider the following claim by Larkin (1977, p.15):

When we read we do more than recognise letters, words and grammatical structures. We also understand; that is, we unpack the words and surface grammatical structures we are presented with into something meaningful, into a deeper structure. A grammar which specifies the relationship between deep and surface structures will specify how surface structures can be unpacked.

Thus, understanding sentences, according to T.G., involves mapping deep structures onto syntactic surface structures. This mapping process is mediated by the operation of transformational rules. According to this idea, when we process a sentence, we de-transform it into its underlying structure. It is suggested that the syntactic complexity of a sentence is proportionate to the number of transformations which have to be undone during processing. This is known as the theory of derivational complexity. Miller and McKean (1964), were among the early proponents of the theory. For them, however, the underlying structure of a sentence was represented by its kernel sentences, the model put forward in 'Syntactic Structures'. In spite of the theory's initial attractiveness, it was found, according

to Fodor and Garrett (1967) that it mispredicted sentences which should be relatively easy or difficult to understand. Moreover, Huggins (1977, p.24) cites an experiment by McMahon (1967) which found that although passive and negative transformations lengthened reaction times, the amount of lengthening produced by a particular transformation did not correlate with the application of other transformations. For example, the increment in reaction times in moving from an active affirmative declarative to an active passive was almost identical to the increment from a negative active to a negative passive. This implies there is some other explanation, (i.e., non-transformational) for the operations required to process the passive construction. We have also seen how the theory of deep structure is at best an inconclusive hypothesis. Transformations themselves are not at all necessary to the processing of sentences by speakers. Fodor and Garrett (1966) conclude that derivational complexity does not correlate in any direct way with understanding and retaining sentences. They suggest that transformational grammars do not correspond to the way sentences are processed. In short, the relation between the alleged native speaker's competence and (psycholinguistic) performance is far less direct than is implied by the theory of derivational complexity. For these reasons, it is suggested here that T.G. does not provide a satisfactory framework in which to analyze syntactic complexity.

It is in some respects ironical that as far as breadth of syntactic description is concerned, T.G. is no more advanced than Bloomfieldian linguistics. The maximal unit of syntactic description in T.G. is still the sentence; no larger unit than this is given formal recognition within the theory. The recent interest in text linguistics, however, has pointed to the existence of supra-sentential discourse structures and has emphasized the structural and lexical ties between individual sentences. The evidence adduced to support this claim is perhaps most succinctly put by Halliday and Hasan (1976, p.1) who argue that

If a speaker of English hears or reads a passage of the language which is more than one sentence in length, he can normally decide without difficulty whether it forms a unified whole or is just a collection of unrelated sentences..... We know as a general rule whether a specimen of our own language constitutes a text or not.

In this sense, 'text' refers to a spoken or written passage which forms a cohesive whole. The goal set for description in text linguistics, therefore, is to explicate the speaker's knowledge of the cohesive devices in his language which make for well-formed texts and hence, account in linguistic terms for their structural and lexical unity. The manifest ability of native speakers to recognise texts has led in text linguistics to the postulation of some structural and lexical level of form to explain this ability. The structural and lexical description of text has an important implication for complexity and difficulty in reading. This is that the successful description of well-formed texts would, definitionally, provide the criteria by which to assess difficult or malformed texts. That is, it would enable the analyst to recognise a text which is not cohesive or where the cohesive ties are imperfectly realised. The latter would be able to predict difficulty in text on the basis of the lack of certain cohesive properties.

Van Dijk (1973) has also pointed to the narrowness of linguistic description within current versions of T.G. He nevertheless incorporates much of the theoretical framework of T.G. into the model he proposes as a possible text grammar. In place of a sentence-generating grammar, he proposes a text-generating grammar. The latter is justified on the basis of stronger generative capacity over sentence-generating grammars. Whereas a Chomskyan-type grammar has as its goal the generation of 'all and only' the well-formed sentences of natural language, a text-generating grammar of the sort proposed by Van Dijk has as its goal the generation of 'all and only' the well-formed texts of natural language. Van Dijk's model accordingly distinguishes, in theoretical terms, between grammatical and ungrammatical texts. In a way reminiscent of Chomsky, the purpose of the text-grammar is to explicate the native speaker's competence for the comprehension of texts. Werlich (1976) also speaks of a text grammar explicating the competence of an ideal decoder and encoder of text. The text grammar put forward by van Dijk is concerned to formalize the relations in text which make for its coherence in this 'ideal' sense. The notion of textual coherence is central to all attempts to produce text grammars. In van Dijk's model (1977a) coherence in text has to do with the logical/grammatical connection between sentences (including reference). It is also concerned with such pragmatic factors as appropriacy conditions, narrative action and, finally, the distribution of given and new information. He considers it possible, in principle, to augment the syntactic component of T.G. with a pragmatic component which delineates the non-linguistic narrative structure of a text.

The purpose of the text grammar proposed by van Dijk is to explicate, firstly, the relations within the sentence and, secondly, relations between sentences, using a process of progressive generalization until finally the global structure of the text is reached. The first stage in the description therefore entails delineating the level of micro-structure in text. Here, the grammar is concerned with the connections between propositions; connections which are expressed through the use of conjunctions, adverbs, etc. Van Dijk (1977b) maintains that there is a systematic correspondence between such connections in text and the logical connectives used in the propositional and predicate calculi. For instance, such connections as conjunction and disjunction (expressed in English by and and or respectively) are paralleled in logical notation by the connections '&' and 'V'.⁶ Similarly, the conditional construction of sentences (expressed in the form of 'if X, then Y') is paralleled in many cases in logic by the material implication connective ' \supset '. The aim, therefore, is for the grammar to be able to specify the inference-relations between propositions in text in truth-functional terms; some inferences will be valid, others invalid (or ungrammatical).

As an example of a valid inference-relation, van Dijk cites the following sentence (1977 a, p.46):

(16) John is a bachelor, so he is not married.

By extension of van Dijk's argument, this inference pattern is captured in the rule of modus ponendo ponens in the predicate calculus:

$$(Vx) (Bx \supset \sim Mx), Ba \vdash \sim Ma$$

- | | | | | |
|-----|-----|-----------------------------|-----|------|
| 1. | (1) | $(Vx) (Bx \supset \sim Mx)$ | A | |
| 2. | (2) | Ba | | A |
| 1. | (3) | $Ba \supset \sim Ma$ | 1 | UE |
| 1,2 | (4) | $\sim Ma$ | 2,3 | MPP. |

Gloss: for every x, if x is a bachelor, then it is not the case that x is married, John (a) is a bachelor; \therefore it is not the case that John is married.

As can be seen, the validity of this inference depends not only on the truth-value of the connectives but also on the 'reference' of the propositions themselves. According to these criteria, example (17) would not be a valid inference:

- (17) John is a bachelor, so Amsterdam is the capital of the Netherlands.

(van Dijk, 1977a, p.46)

It is not necessary that the propositions are explicitly related by connectives, only that the reference between propositions centres on a common topic for the completion of a valid pattern of inference. The latter also includes, therefore, the pragmatic notion of presupposition in sentences. This micro-level description, however, represents only the first stage in the description by the text grammar: the description of all and only the valid inference-patterns between successfully-referring propositions in text. Such a formal semantic description of text in van Dijk's model is based on the use of propositional, predicate and modal logics as the descriptive apparatus. This stage in the analysis is primary because "... the study of relations between sentences in a discourse will have first of all to show how the meaning and reference of sequences of sentences depends on the meaning of their component sentences" (1977 a, p.44). Moreover, van Dijk claims that many of the relations obtaining between clauses in a compound sentence also obtain between sentences in sequence, and conversely.

The next stage in the analysis by the grammar is to describe the macro-structures of the text. Macro-structures are compound topics, which are the result of successive amalgamations of the smaller topics of the sentences in text. That is, each sentence may make specific information-contributions so that a composite topic of discourse is generated, a macro-structure. Such macro-structures themselves make information-contributions to generate still larger macro-structures until, finally, the global macro-structure is generated. In short, the relation between micro-and macro-structure consists of a sequence of successive approximations from the local topics to the larger global topic. These relations are seen as hierarchically-organized and illustrate a process of progressive generality in the description of the text. The information-structure of the text is also taken to constitute the narrative schema, or action provided by the plot. Thus, coherence in text according to von Dijk's model is a function of the degree of interrelation between these various levels.

The text-generating grammar as proposed in von Dijk (1977 a) is, however, programmatic. More specifically, it is proposed as an outline or set of methodological principles for the construction of a text-generating grammar. Such a grammar would include (as well as a sentence-generating component), "a pragmatic component, a reference semantics, a semantics with world-knowledge

interpretation conditions, and a macro-semantics" (p.7). The function of the grammar is to explicate, at a high level of generality, the process whereby macro (deep) structures are mapped on to the syntactic surface structures of the text. Van Dijk admits to some extent the great magnitude of this task when he says (p.8) "We do not want to specify the precise structure of such a grammar but only to give some of its possible FRAGMENTS, and some relationships between the semantic and pragmatic fragments of such a grammar."

Owing to the fact that it is based on T.G. theory, the text-generating grammar of van Dijk is bedevilled by the same methodological and theoretical problems. For example, as van Dijk notes, the generation of the content of the information-structure of a text is a 'creative' process and has to do with such factors as world-knowledge, individual biography, etc. If we concede that this is a truly creative process, then it is difficult to see how texts can be generated from a finite set of rules in a text-generating grammar. It is certainly not easy to envisage the form of a component in the grammar of "world-knowledge interpretation conditions." This is analogous to the problems in T.G. (see Chomsky, 1965), where the grammar attempts to specify selection restrictions by sub-categorization restrictions in the base component. The intention here, for example, was for the grammar to distinguish between semantically acceptable and semantically anomalous sentences. This attempt failed precisely because of the creative use of language. A sentence which is semantically anomalous at one period in time may, conceivably, be semantically acceptable at another because of changes in the state of world-knowledge. Moreover, the distinction between acceptable and anomalous sentences is more of a gradient property than a 'yes' or 'no' one; some sentences are more/less acceptable than others. This contention applies particularly to the text grammar's attempt to differentiate between valid and invalid inference-patterns. Some inference-patterns will be more/less valid than others. This is realised to some extent by van Dijk, who cites the following sentence as a 'borderline' pattern of inference:

(18) John is a bachelor, so he buys too many records.

(1977 a, p.46)

The validity of this inference is certainly conceivable in certain contexts or states of world-knowledge. Thus, any decision-procedure for the validity/invalidity of inference-patterns between propositions must account for their

creative (contextually-determined) properties. It is far from clear, however, how this is best achieved in the model proposed by van Dijk. By the same token, the distinction between well-formed and malformed texts is more of a gradient property than a clear-cut one. Some texts will be more grammatical than others, according to the extent of the validity of the inference-patterns which they display and coherence in the distribution of given and new information between sentences.

In conclusion, the linguistic procedures posited by van Dijk for the analysis of text seem eminently feasible as descriptions of text. But the claim to generate well-formed texts seems much less certain and remains to be seen. It would, also, surely be an advantage in any text grammar if the cumbersome logical machinery envisaged by van Dijk could be replaced by descriptive procedures which use the connective items in natural language to 'name' the connections. There is an obvious circularity in the translation of the connections in natural language into those of another (artificial) language. There seems a good case to be made for using the natural meta-language of English to talk about connections between sentences (see chapter four of the present thesis). Perhaps the greatest practical limitation of van Dijk's model, however, is that it is really dealing with the competence of the encoder/decoder of texts rather than with aspects of performance. From the point of view of providing a practical set of procedures for analyzing complexity or difficulty in text, van Dijk's model seems very limited in scope.

The notion of information-distribution in text and how it is manifested structurally has led linguists and psychologists to postulate less formal descriptive procedures than van Dijk. The intention here is to delineate the ways sentences are related by virtue of a commonality of reference. As for van Dijk, this 'commonality of reference' constitutes the various topics which are responsible for the cohesion in text. In this account, cohesion is not restricted to the grammatical devices of connection but includes lexical devices as well. Cohesion in this sense is therefore also a qualitative concept: the complexity or difficulty of a text is a function of the extent of cohesion it displays. More precisely, it depends on the way this cohesion is organised; how larger topics are composed of smaller topics; the part of a sentence occupied by a topic and its relation to other parts of the sentences etc.

Thus Grimes (1975) analyzes text into three principal components: content, cohesion and staging. Two types of framework are proposed to represent the

structure of a text. The first is a graph, structured in the form of a hierarchy, which is composed of categories derived from case grammar supplemented by rhetorical predicates. The second framework is in the form of a table which is used to display the arrangement of the information in a text (e.g., the actions, settings, participants, etc.). In Grimes (1978), topics are analyzed according to their thematic importance. According to the latter, topics are graded to one of a series of levels of importance in the text. This constitutes the staging in text, which "reflects how the speaker calibrates the importance of different parts of what he himself intends to say" (p.104). For Grimes, a topic is the referential common-ground between speaker and hearer "about the identity of certain objects in the real world. It may also be agreement about certain events or about certain relations that hold." (p.104).

In the course of a text the topic, once established, may be developed, according to Grimes, in at least three different ways. Firstly, it may be expanded, where things are added to the referential case. Secondly, it may undergo shifting, where some referents are omitted and others added, thus changing the common ground of reference. Finally, it may be split. Here, global topics are divided into a higher level and a lower level topic. Meyer (1975) and Clements (1976) have refined this notion of topic-splitting into a higher and lower level. The framework they propose is a hierarchy which shows the global topic, the local topics discussed in relation to the global topic and whatever lower level topics are discussed in relation to the local topics. This analysis produces a topic tree of indefinite depth. For Clements, three rules operate in the analysis for the global-local topic hierarchy:

- (i) Topic rule: identify the topic of each clause or simple sentence.
- (ii) Old/new rule: decide whether the topic is new (i.e., not mentioned before) or old (i.e., mentioned in an earlier topic or comment). If the topic is new, then assign it one level below the previous topic. If it is old, then assign it the same level as its first mention.
- (iii) Co-ordinator rule: if a topic is co-ordinated with an earlier topic or comment, assign it the same level as the earlier topic or comment.

These proposals of Meyer and Clements strongly suggest that there may be a hierarchy of topics in a text. One of the psychological implications of the latter is that a low position in the hierarchy makes for worse recall of a topic than one in a higher position. In short, this work addresses itself

to the question of what factors in text affect text-learning and recall.

Crothers (1975) also refers to the distinction between 'old' and 'new' information in text structure but refers, in addition, to a third type of information, namely, contrastive. The latter is defined as a term "which is viewed as a textually new member of a textually old semantic category, the category having already been implied by the term against which the present one is being contrasted." (p.23). Such a contrastive relationship is illustrated by example (19) below:

- (19) A nation's freedom conduces so much more to art and creativity than a nation's dependence.

In example (19) there is a contrast between freedom and dependence, although both terms belong to the semantic category of nation. Crothers is concerned to develop a descriptive procedure for the analysis of paragraph structure. To this end, his analysis attempts to derive the textual structure of the paragraph in terms of the semantic representation underlying the sentences and the inference-relations that obtain between the constituent propositions. As for van Dijk, Crother's analysis deals primarily with propositional connectives between underlying propositions. Firstly, the sentences are analyzed into their various propositions. Secondly, the various connections between the propositions are examined. Finally, the information-structure of the whole paragraph is analyzed by means of the triple categories of old/new/contrastive information. Each paragraph analysis is organized in the form of a tabular array, showing the content of the passage.

In sum, the analyses of van Dijk, Grimes, Clements, Meyer and Crothers are linked by a common pedagogical implication: successful comprehension of text depends not just on 'what' is said but the 'way' it is said. This suggests that the organisation of prose-text can enhance or impair comprehension and recall of its content. Moreover, these approaches show that cohesion between sentences is as great a determinant of comprehension in text as the organisation of the individual sentence, if not more so. Given this, the analysis of T.G., with its confinement to the individual sentence does not provide a satisfactory framework in which to study complexity in text. With the exception of van Dijk, these approaches have also shown that it is not necessary (or even desirable) to invoke the notion of deep structure of text. Accordingly, complexity and difficulty in text is a function of the distribution of of information throughout the text, the connections between the various

propositions or clauses, and the referential common ground between author and reader. All these analyses approach the text via its semantic organisation. It is argued in the present thesis, however, that syntactic complexity and difficulty in reading is also a function of the syntactic organisation of sentences and the relations between them. A proper examination of this question entails providing possible syntactic categories with which to approach the syntactic complexity of text. The present thesis addresses itself specifically to this problem.

SCOPE AND APPROACH OF THE PRESENT WORK

The main contention of the present thesis is that in spite of the varied and detailed approaches to complexity in text, we still do not possess adequate grammatical criteria by which to assess complexity. Moreover, as we saw in T.G., the categories proposed are based on the analysis of competence, not performance. Thus, not only do we need the necessary descriptive categories; we also need categories which are based on actual texts, as aspects of performance. Only in this way can we hope to develop a realistic and accurate analysis of complexity and difficulty in text. We have also seen that many of the various approaches we have considered are principally concerned with (semantic) relationships between sentences; consequently, little attention has been given to the relationships in the sentence itself. 'Difficulty' in reading, it is argued here, is also a function of syntactic complexity within the individual sentence, not just between sentences.

Before we can successfully analyse complexity between sentences we must have some idea of syntactic complexity 'within' the sentence. The latter entails that we have the means to delimit the boundaries of the sentence or clause. The reason for the paucity of adequate definitions of the sentence, according to Lyons (1977, p.629), is that "Chomsky and his followers ... have been content to operate with the assumption that native speakers have an intuitive appreciation of the fact that certain strings of forms are sentences and others are not. But they have failed to give any account, even in principle, of the way the sentence as a theoretical construct within the linguist's model of the language system is related to the sentence as a contextualized product of language - behaviour." It is hoped that the definition of the sentences in this thesis satisfies this requirement of being a "contextualized product of language-behaviour."

Definitionally, the term 'sentence' is used in this thesis to refer to the communicative function of the independent declarative clause; that is, the communicative function of presenting new information. But the term 'clause' is used to refer to the grammatical structure of the sentence. The grammatical structure of the sentence consists of the following constituent categories:

- (20) Subject, Verb, (Object), (Complement), (Adjunct).

The brackets indicate the optional nature of the items they enclose. Of course, the orthographic sentence frequently consists of more than one independent (declarative) clause but this is, strictly speaking, a compound structure. The clause as defined in example (20) represents the 'minimal' grammatical form of the sentence, which may be expanded by other clauses. Syntactic complexity within the sentence, therefore, has to do with the intrusion of extraneous grammatical structures within and between the constituents shown in example (20). In addition, it has to do with the delaying of parts of the structure of the clause by expansion of other constituents. In short, complexity in the sentence is a function of the degree of interruption and delaying of the structure of the clause. Sentence complexity is intimately bound up with the internal grammatical structure of the sentence (Bram, 1978). This is not, however, the only approach to complexity in the sentence in terms of its surface syntactic form.

The most useful and widely-used approaches to syntactic complexity within the sentence are readability formulae and cloze procedure. In the case of readability formulae, the readability level of a given sentence is held to be a function of the number of words it contains, the frequency of the individual words, the number of subordinate clauses or prepositional phrases, the proportion of concrete as opposed to abstract words, etc. However, "Formulas measure only one aspect of writing-style ... Formulas do not touch on organisation, word order or imagery in writing ..." (Klare, 1963, p.24; my underlining). On the basis of these criteria, the difficulty of a passage is ascertained from a set of statistical descriptions. Cloze procedure, on the other hand, provides for an experimental procedure to measure the relative importance of individual words in the text. It does this by blanking out certain words in the text, which students are then asked to complete. Students' error patterns are taken as indicators of the relative simplicity/difficulty of the words in question. The distinct advantage of readability formulae is that they provide simple, automated procedures which can be applied to large amounts of prose-material to test readability level. The advantages which this affords to editors who deal in large amounts of prose is that both cost and human time are reduced. Nevertheless, readability formulae do not show 'why' a given sentence is complex. As Huggins (1977) argues:

Readability measures ... present statistical rather than structural descriptions of complexity. Unfortunately, correlation does not imply [prove] causality, and although passages that are highly readable tend to have

short sentences, and few prepositional phrases or clauses per sentence, it does not follow that writing short sentences with few prepositional phrases yields highly readable text... Thus, although sentence length correlates with syntactic complexity, it cannot be used to explain it.

(p.19)

It is possible to have sentences which are abnormally long but structurally quite simple and therefore easy to read. At the other extreme, there are sentences which are quite short but whose syntactic structure renders them quite difficult to read at first (Glazerfield, 1970-71, p.12-13). As an example of a sentence which, although abnormally long, has quite a simple internal grammatical structure, consider example (21) below.

- (21) S V O
- The museum contains [various specimens of tropical fish,] [the
skeletal remains of dinosaurs from the prehistoric age] [a
variety of butterflies and insects of all shapes and sizes,]
[authentic models of African and Indian elephants and other
wild animals,] [an array of the fossilized remains of plants
and small fish], [a very impressive collection of meteorites
and rock forms;] and, finally, [a display of military equipment
and weaponry and the latest developments in lazer technology.]

Although example (21) contains seventy-six words, it is not a difficult sentence to read. The reason is that the sentence's grammatical structure is quite simple, consisting of a subject, transitive verb and seven object nominal groups. This simple transitivity structure creates in the reader a strong expectation of what is to follow in the sentence (i.e., an object or objects). It is in cases where the syntax of the sentence impairs this reader expectation that syntactic complexity and difficulty may occur in text. Sentence (21) is to be contrasted with the less usual word-order of example (22):

- (22) PP PP V S
- [Down the street] and [around the corner] ran Louis.

The superscripts in this example denote two prepositional phrases, an intransitive verb and a subject. The syntactic order of the sentence thus

is subordinate to the need for syntactic clarity. When syntactic clarity is threatened, part of the structure of the(long) subject must be delayed, even at the expense of reducing its markedness.

The first section of chapter two examines the internal structure of the major constituents of the clause referred to in chapter one: the subject, verb, object, complement and adjunct. It is shown that the structure of these items consists of a head which is preceded and followed by modifying structures (auxiliary structures). Thus, in the case of the verb phrase, there is a main verb and one or more possible auxiliary verbs. Similarly, the structure of the subject, object, complement and adjunct consists of a noun head which may be premodified and postmodified, or both. The noun head forms the central semantic core of these latter structures. Moreover, the reader relies on such heads for his basic semantic interpretation of the sentence. Syntactic clarity therefore requires that the relations between the heads of the various constituents are not obscured by excessive premodification and/or postmodification. In a similar way to the long subject, excessive modification functions to delay the other noun heads in the sentence and hence, the sentence boundary itself. This section proposes an intuitive limit to the amount of modification which may precede and follow the head in the subject, object, adjunct or complement. The second section of chapter two develops this notion of headed structures in relation to apposition within the clause. Apposition concerns a syntactic relation between 'like' grammatical structures; more specifically, it is a relation between headed structures where the second headed structure further specifies the meaning of the first. The second modifying structure in the apposition relation, in most cases, is thus an optional part of clause structure. It is not directly integrated into the structure of the clause but interrupts the relation between its major constituents. Apposition can also interrupt the structure of the major constituents themselves. In sum, the suggestion of the second section of chapter two is that appositional structures should not be used excessively within the sentence. If this rule is not observed, then the repeated interruption of the structure of the clause will make the predictive task facing the reader all the more difficult.

Chapter three follows on directly from chapter two insofar as it, too, is concerned with interruption of the structure of the clause by extraneous modifying structures. This chapter considers the evaluative and subjective aspect of text and how this may be used to comment on or assess any item in the clause. This aspect of modification is considered under the title of Interpolation. In terms of its mobility within the clause, interpolation is

shown to be a special extension of the adjunct function. Interpolation may vary from a nominal group to a whole clause or sequence of clauses. It is at the level of the clause that interpolation often entails the most serious consequences for complexity within the clause. More precisely, when the interpolation is in the form of an independent declarative clause and interrupts the structure of another independent declarative clause, there is a serious breach of the communicative function of the interrupted clause. That is, the presentation of new information in the interrupted clause is confused by that in the interpolated clause. Thus, although interpolation in the form of a declarative clause is a perfectly acceptable form of interruption, there are strict limits on its use. Chapter three also includes a discussion of the treatment of interpolation within T.G. It is suggested that the communicative function of interpolation, in the form of an independent declarative clause, is confused in T.G. with that of the non-restrictive relative clause. Consequently, transformations in this sense are not sufficiently sensitive to the contextual functions of the clause.

Chapter four considers the second dimension of syntactic complexity in text under the title of Clause Relations. Here, the categories developed in chapters one to three (i.e., the long subject, the nominal group, the appositional group and interpolation) are considered in terms of their delaying and interruptive function within the framework provided by Winter's theory of clause relations (Winter, 1977). In this chapter, it is argued that special connective items in discourse function to connect members (of no fixed maximum length) in a binary relationship. Such clause relations have no necessary relation to the orthographical divisions in text. The second type of syntactic complexity is concerned with how the four categories that are considered can impair the cohesion in text by delaying and interrupting crucial clause relations. The conclusion, chapter five, represents an attempt to synthesize the various categories of complexity that are considered into a coherent framework.

Clearly, the principal point of reference for both dimensions of complexity considered in this thesis is the independent (declarative) clause. A definition of the clause is proposed in example (20) above. This presupposes, of course, that the clause represents a valid unit of perceptual segmentation of text for the reader. This is a crucial supposition. What evidence can be adduced in support of it? Perhaps the most important experimental studies in this area are the 'click' experiments conducted by Fodor, Bever and Garret,

(Fodor and Bever, 1965; Garrett, 1965; Garrett, Bever and Fodor, 1966). The purpose of these experiments was to test the 'psychological reality' of (surface) constituency structure. It was found that in sentences presented to subjects which had superimposed displaced (extraneous) click sounds, the clicks in verbatim recall were attracted towards major constituent boundaries in the clause. Most importantly, however, these experiments also revealed a marked tendency for the clicks to be attracted to clause boundaries. The latter suggests that subjects were too concerned with processing the clause itself when they heard the displaced click; consequently, they wrongly identified it to have occurred at a clause boundary, where they momentarily pause in processing. We can also infer from this that the clause represents the unit by which sentences are processed and organized in memory. Further support for the integrity of the clause as a surface constituent is provided by Freedle (1978). This study showed that subjects took longer to complete ambiguous sentence fragments as opposed to control fragments, when the fragments were incomplete clauses. However, subjects in the study displayed no such difference when the fragments they were presented with were complete clauses. Freedle concludes that the clause represents a valid unit of segmentation in the processing of text. Finally, Flores d'Arcais (1978) also claims that the clause represents an important level of perceptual segmentation:

The unit within which syntactic processing would take place in order to extract semantic interpretation would normally correspond to the clause. Once a speech segment has been perceptually isolated as a clause, processed and understood, its information is stored in memory and processing of the next segment can take place.

(p. 156)

If the clause is indeed such an important unit of perceptual segmentation in text, as the studies above suggest, then the interruption and delaying of the structure of the clause is all the more serious. This means that the excessive use of the categories that are postulated in this thesis functions to disorientate the reader, making his predictive task more difficult. This has implications for the reader's assimilation of the meaning of the sentence and his comprehension of its content. Equally important, the interruption and delaying of the structure of the clause has the effect of interrupting and delaying the still larger

unit of perceptual segmentation in the form of a clause relation. All the categories for analyzing complexity which are put forward are, of course, hypotheses which must be psychologically validated. Nevertheless, they are supported as meta-theories for complexity by the writer's own intuitive appreciation of the difficulties involved in reading.

The data chosen for the development of the categories of apposition and interpolation is formed by extracts from Open University course units and current newspapers and periodicals. Three social science course units were randomly selected for detailed examination. These were: D305 Block 8, (14-15); D305 Block 9, (16-17); and D301, (13-15). It is hoped that a possible narrowness of empirical scope of these units is supplemented by the non-academic material also examined. Finally, a glossary of some of the technical terms used in this work is provided at the end of the thesis. For convenience of reference, the terms given are listed alphabetically. It is hoped that by providing definitions even of terms extensively examined in the main body of the work, the reader will have a set of useful mnemonics which he can refer to at need.

Chapter One: The Long Subject Construction

SECTION 1.1.

As readers of texts we have a dual need for clarity of syntax and emphasis. Broadly, this means that we will be motivated to read a text only if the text displays, in the first place, a clear syntax; that is, the main verb of the sentence must not be delayed too long or this will hinder our comprehension of the sentence. In the second place, we are persuaded to read a text successfully by assuming that the writer emphasizes (marks) those sentences or parts of sentences that he considers important. This assumption entails that every sentence is 'unimportant' (unmarked) unless it is signalled to be important by the writer. These two complementary needs, however, are not of equal importance to the process of reading; rather, successful emphasis is a function of clarity of syntax.

What do we mean exactly by clarity of syntax and emphasis? Clarity of syntax is here used to refer to the common sense notion of using language in texts to communicate to a reader in as clear a way as possible; more specifically, how to transmit a message without causing confusion or creating an obstruction to his assimilation of the sentence's meaning. The latter takes into account the well-known difficulty of excessive sentence-length; it also includes the point that a succession of long sentences becomes intolerable to the reader. In contrast, emphasis has to do with the practice of marking certain parts of the clause to signal their relative importance to the text as a whole. Here, certain material in the clause is highlighted by its syntactic position.

We may consider first of all the question of emphasis and the syntactic devices used to achieve this. It is well-known that in English it is normal to present the subject of the sentence first, with the verb following. This is the usual sequence so long as the subject does not exceed a certain length or degree of complexity. For practical purposes, however, we may assume a standard of simplicity. Thus, the shortest (and hence, most simple) subject may be a pronoun, as underlined in the sentence, 'This is true'; alternatively, it may be a noun phrase consisting of the sequence determiner + noun, as in 'The man'. The noun phrase may vary from two words, as in the last example, to four words (as in 'The kindly old man')

or even more.

Taking this definition of the usual (short) subject as our datum, the long subject construction has two definitions. Firstly, the long subject may expand the subject noun phrase. Here, the noun phrase is further specified by the addition of relative clauses and their accompanying structure. The accompanying structure of the relative clause may itself contain verbs which precede the main verb of the sentence. In the example given below the underlined items are the verbs contained by the clause, while the relative clauses themselves are denoted by round brackets:

The kindly old man (whom you described so graphically and whom you thought a complete unknown) is none other than the President.

In its second definition, the long subject begins with that (more generally, with a WH-pronoun, denoted by such words as how, what, whether, etc.), which is used to embed a further clause, followed in turn by the verb be and, optionally, by a complement (denoted in the example below by brackets):

That - clause V (C)

Alternatively, the long subject may take the form of an embedded non-finite clause, followed by the auxiliary and complement. In the examples given below, the two subordinated that-clauses and infinitival clause function as the long subject (underlined in the examples):

- (1) That small children should be allowed to roam, after dark, and to knock on doors in neighbourhoods some of which have a dubious reputation, never fails to astound me.

(Observer letter, 18/12/77)

- (2) Why he wants just that formula is baffling to me.

- (3) To argue that Jesus himself would welcome this interest in the shroud because "he rejected the temptation to give physical miraculous tokens of his Messiahship knowing impressions so created could not aid his purpose" is somewhat contrary to the evidence.

(Guardian letter, 15/5/79)

Consider the following examples of the long subject, which are given in order of subject-length and complexity:¹

- (4) (The question whether their techniques could aid the police in tracing a wanted person) immediately^A arises^V.
- (5) (Courses which do an all-round job and which provide the student with very powerful motivation for learning) exist^V.
- (6) (The day when it will not be worth making a statement about style in painting or literature unless the statement can be so precisely formulated that it can be translated into a computer programme) is^{vi} coming^{vii}.

The long subject in all these examples is denoted by round brackets, while the verbs and adjunct are indicated by the superscripts. Firstly, we must note that all three examples are marked. By being placed in initial position in the sentence, with the verb(s) and adjunct following, the long subject draws attention to itself as the most important item in the sentence. In this way, the long subject is frequently successful as a device for achieving emphasis in the sentence. Example (4) has the structure S A V, while example (5) has S V and example (6) has S V; (i.e., subject followed by auxiliary verb) Vii (i.e., the main verb). Also, the main verb which follows the long subject in each example is intransitive. Using our earlier notion of clarity of syntax, it can be seen that the degree of sentence difficulty gradually increases through examples (4) to (6). We find it progressively more difficult to assimilate the meaning of the sentence as we read through this sequence. The reason for this increase of difficulty is best seen in the extent of the subordination contained by the subject which precedes the main verb in each example. Example (4), for instance, contains only one layer of subordination in the (long) subject (signalled by which). Because only a single subordinate clause precedes the main verb arises, the long subject can be tolerated by the reader; the main verb is not delayed too long after the subject and the sentence can be read with relative ease. This latter point may, in fact, be put forward as an hypothesis about our comprehension of sentences; that is, that the main verb is the most crucial item for comprehension of the sentence as a whole. Fodor (1971), Schank (1972), and Winograd (1972), have all emphasized the centrality of the verb in the interpretation of sentences. Until we read the main verb of the

sentence, we simply don't know what the sentence is about. The perceptual task facing the reader in the case of a complicated sentence, for example, is to identify the subject and main verb and relate the two successfully. This crucial cognitive process is impaired if the main verb is delayed too long in the sentence. The primacy of the main verb in sentence comprehension is illustrated by the following incomplete construction:

The kindly old man whom you described so graphically
in your talk yesterday

We cannot understand this sentence because the main verb is missing. By the same token, when the main verb of a sentence is delayed by too many levels of subordination, our comprehension is likewise delayed. This delaying of the reader's comprehension is particularly characteristic of constructions with long nominal clauses in subject position, as illustrated by examples (5) and (6). In fact, nominals in subject position have been rated as more difficult to comprehend than nominals in object position (Bever and Weksel, 1966).

Example (5) is less easy to understand than example (4) because it contains two subordinate clauses before the main verb. (Signalled by two occurrences of which). Even more extreme in this respect is example (6), which contains three subordinate clauses (signalled by when, unless and that, respectively); consequently, it is the most difficult of all to understand. As in example (5), the auxiliary verb is and main verb coming are placed in final position in example (6). It is interesting to compare examples (4) to (6) with a construction cited in Clark (1977, p. 65):

--S--

- (7) (The man offered a hundred dollars for the bottle of 1962 vintage Mouton Rothschild) ^Vis ^Cmy uncle.

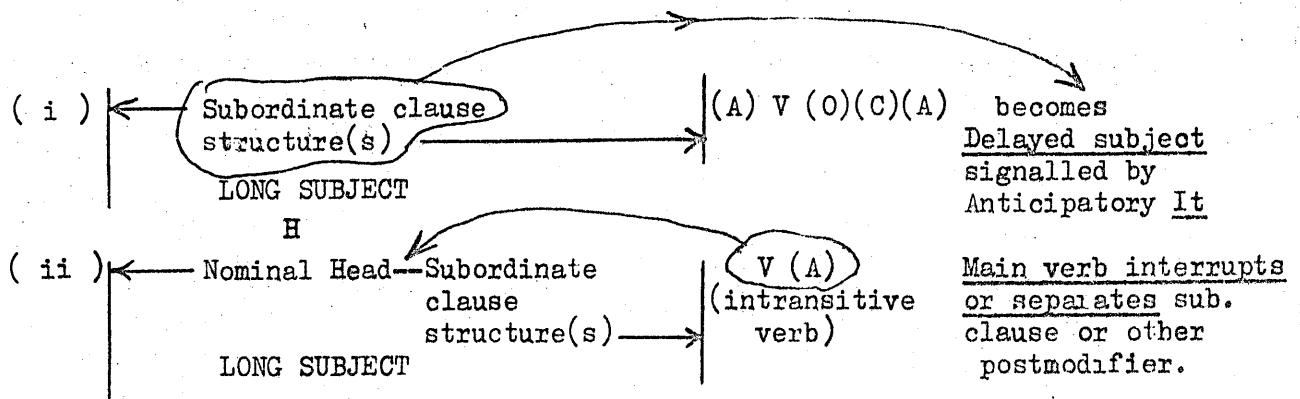
Example (7) contains a long subject, which is shown in round brackets. Clark does not attribute the relative difficulty of this sentence to its long subject. Although he maintains that the sentence places a more than average load on the reader's memory, he attributes this simply to the deletion of the relative pronoun and past form of the copulative verb between the man and offered. The length of the sentence, he argues, makes it difficult for the reader to infer that this deletion has taken

place. But this seems only a partial explanation for the difficulty of example (7). Such difficulty is more properly seen as a function of the long subject before the main verb is. As further proof that too much subordination in the subject delays comprehension of the whole sentence (by delaying the main verb), consider example (8) below:

- (8) ^H [The explanation ((advanced by Phillip J. Klass in an article in Aviation Week (Vol.85 No.8), in which corona discharges, in the shape of balls of gas thrown up by the power lines, are held responsible for the UFOs themselves))] ^V ^A ^C is more convincing, however. ^A

In this example, the superscripts denote the noun head, main verb, complement and adjuncts. The double round brackets denote the post-modifying sub-clauses of the noun head, while the square brackets indicate the total length of the subject. Example (8) is a classic instance of attention being drawn to the subject of the sentence at the cost of reducing its clarity of syntax. Because emphasis in the sentence is subordinate to clarity of syntax, however, unacceptable sentences like example (8) must be rewritten, even though this may entail reducing the markedness of the subject.

As is fairly well-known, English Grammar has two main ways of improving the comprehension of excessively long subjects. The first is where we invert the sentence so that the long subject comes last (See Section 1.3). The second is where we delay or interrupt the subject according to whether it is a subordinate clause or whether it is a nominal group having a subordinate clause. (This is variously discussed as 'discontinuous noun phrases' in Quirk and Greenbaum and others). These latter two ways of delaying or interrupting the subject are made clear by the following diagrams:



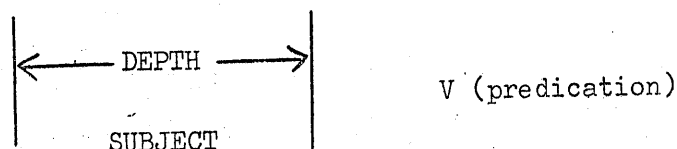
In both (i) and (ii), the analysis is $S_I V S_2$, but in (i), S_I is the item It signalling real S_2 to follow, and in (ii), S_I is nominal head of

real S, so that the postmodifying structure of real S follows as S_2 , as can be seen below if we rewrite examples (4) to (6) in the following way:

- (9) [S_1 The question] A immediately V arises [S_2 whether these techniques could aid the police in tracing a wanted person.]
- (10) [S_1 Courses] V exist [S_2 which do an all-round job and which provide the student with very powerful motivation for learning.]
- (11) [S_1 The day] Vi is Vii coming [S_2 when it will not be worth making a statement about style in painting or literature unless the statement can be so precisely formulated that it can be translated into a computer programme.]

In examples (9) to (11) the long subject (which is now in a delayed, discontinuous sequence) is shown in square brackets. The delayed subordinate structure (which occurs after the main verb) is indicated by single lines. The superscripts S_1 and S_2 denote the real subject and expansion of the subject, respectively; V_i denotes the auxiliary verb and V_{ii} the main verb. In example (9) we have interrupted the long subject by the adjunct and main verb, both of these being inserted between the noun phrase The question (determiner + noun head) and the relative pronoun whether with its clause. In example (10), we have inserted the main verb exist between the plural noun courses and the relative pronoun which and its clause. Finally, in example (11), we have inserted the auxiliary verb be and its lexical head (present participle) between the noun phrase The day and its relative pronoun when. The result of all these changes is to provide an increased amount of sentence clarity. Examples (9) to (11) are not totally unmarked, however, in spite of these changes. Some markedness now comes from the front position of the main verb, as in example (10), which begins "Courses exist which..."

In sum, all the examples of the long subject discussed so far point to sentence depth as a potential cause of syntactic complexity in sentences. Sentence depth is here taken, broadly, to refer to the extent of the expansion of the simple subject preceding the main verb of the sentence. This can be illustrated in the following way:



We have taken the sentence depth of the long subject as proportionate to the layers of subordination preceding the main verb.² It is suggested that the reader can tolerate more depth if it occurs in a discontinuous sequence; that is, depth which occurs after the main verb. The notion of depth is given a much more formal and detailed specification within Chomskyan linguistics. Depth is a crucial part of the argument put forward in T.G. for the grammatical creativity of a language such as English. T.G. theorists, however, are concerned with the characterisation of the native speaker's competence and, as such, regard as irrelevant all real-world performance-constraints on the use of language (Chomsky, 1965, p.3). We have posited the need for clarity of syntax in relation to the long subject construction as a performance-constraint. The latter is a factor (no doubt having important psychological correlates) which actively controls the way we speak and write. Because the notion of depth is so central to our analysis of the long subject construction, we will devote the following section to a more thorough treatment of this concept. Special attention will be given to the treatment of depth within the formal framework of Chomskyan linguistic theory and the implications which the present approach has for this.

SECTION 1.2

The advent of Chomskyan linguistics (1957) represented an attempt to mathematize the formal basis of language study. Formal grammatical description became more abstract: the construction of a grammar, or particular classes of grammars, was part of a more inclusive grammar-type, or theory of grammars. Special classes of grammars were used (called phrase structure grammars) to generate classes of formal languages. By generating such languages (or stringsets) it was hoped to delimit the boundaries for formal languages and, by implication, those of natural language. The aim of this approach was to develop some finite means (a grammar) for specifying all and only the acceptable strings (sentences) of a particular language, L. The assumption was that human languages are infinite classes of stringsets; through such processes as conjoining and embedding, it was argued, there is theoretically no limit to the range of sentences which speakers can produce. However, it was also argued that speakers can nevertheless differentiate between grammatical and ungrammatical sentences in their native language. Since the creative potential of language is so infinite, speakers must have at their disposal some finite specification of what constitutes (and does not constitute) a sentence in their language. In short, the speaker is able to 'decide' which are grammatical and which are ungrammatical strings in his language. The latter property makes the class of natural languages recursively-enumerable languages. The finite set of rules needed to produce the sentences of natural language was held to reside in the speaker's internalised grammar of his language: his linguistic competence (Chomsky, 1965).

Phrase structure grammars are such finite devices for specifying languages. These grammars delimit sets of terminal strings and assign to each a particular phrase structure. The rules of the grammar operate by replacing one non-terminal symbol in some finite vocabulary, V, by a corresponding terminal symbol or set of such symbols. The rules of the grammar are rewrite rules which operate from left to right. As an illustration of this we may consider the class of context-free grammars, which generate context-free languages. Suppose, for example, we wish to generate a (formal) language which contains strings of the following form: ab, abc, abcdcd, abcdcdcd.. That is, we have to provide a finite number of phrase structure rules to generate a right-branching language: a language which creates new strings

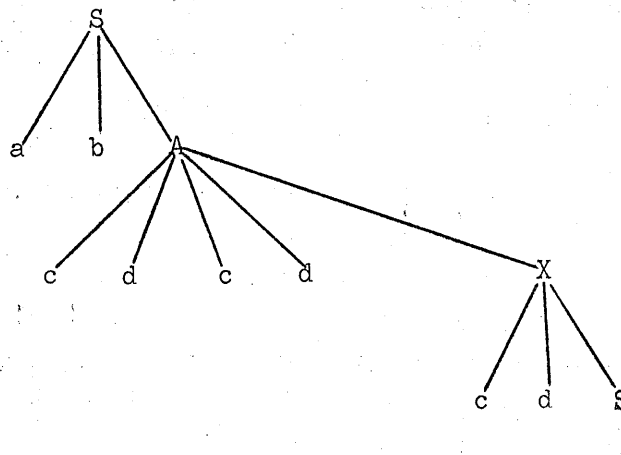
at the right-most end of a derivation. The rules of this grammar must be right-branching or right-linear. They must also be recursive: that is, capable of indefinite extension. A right-linear rule takes the following general form:

$$S \longrightarrow a X$$

Each application of this rule will generate one terminal symbol, followed by a non-terminal symbol that may be further rewritten. Thus, a right-branching grammar needed to generate the language just given would consist of the following set of rules (where each bracket indicates a possible option of choice):

$$\begin{aligned} S &\longrightarrow a b (A) \\ A &\longrightarrow c (d c d (X)) \\ X &\longrightarrow c d S \end{aligned}$$

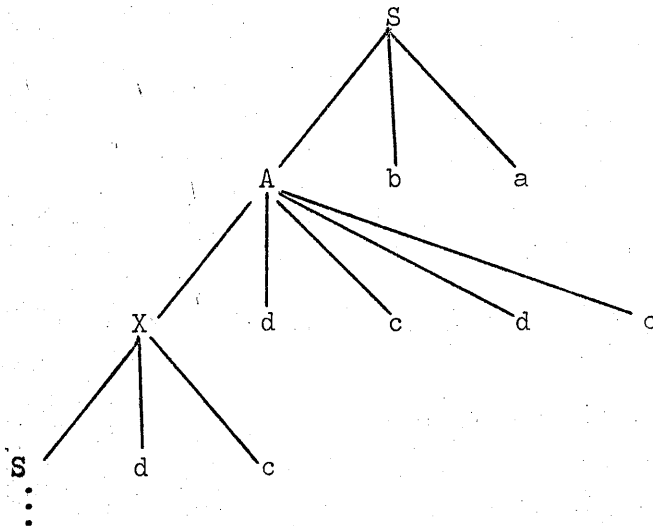
When applied, these rules would produce the following phrase marker:



A left branching language which corresponds to this would have such strings as ba, cba, dcdcba, dcdcdcba.. Thus, creativity in this language is provided at the left-most part of the derivation. A left-branching grammar for this language would be:

$$\begin{aligned} S &\longrightarrow (A) ba \\ A &\longrightarrow ((X) d c d) c \\ X &\longrightarrow (S) dc \end{aligned}$$

These rules would generate a phrase marker such as the following:



A sentence from a natural language such as English which corresponds to the right-branching grammar just given is shown below (where the constituency-structure is denoted by numbered square brackets):

¹
 [This is ²
 [the dog belonging to ³
 [the teacher of ⁴
 [the
 daughter of ⁵
 [the gardener of ⁶
 [the friend of ⁷
 [John. ⁷ ⁶ ⁵ ⁴ ³ ² ¹

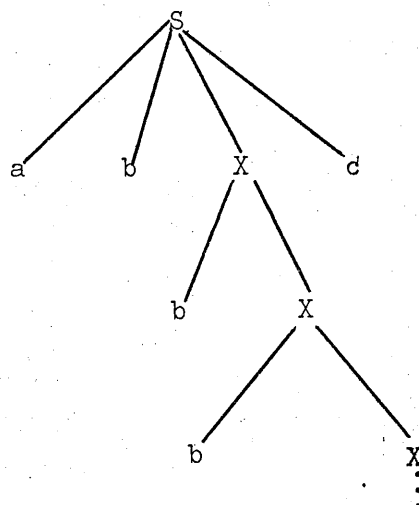
The left-branching equivalent noun phrase of this sentence is:

¹ ² ³ ⁴ ⁵ ⁶ ⁶ ⁵ ⁴ ³
 [[[[[[John] 's ₂ friend] 's ₁ gardener] 's daughter] 's
 teacher] 's dog.] ₃

For the sake of completeness, mention must also be made of so-called self-embedded constructions. Unlike the left- and right-branching structures just illustrated, creativity here occurs within the centre of a derivation. A formal language corresponding to this would be of the form $a b^n c$ and would contain strings such as abc , $abbc$, $abbbc$... A grammar to generate this language is specified by the following rules:

$S \longrightarrow a b (X) c$
 $X \longrightarrow b (X)$

which produce a tree such as:



A self-embedded construction from English which would correspond to the output of this grammar would be:

¹[The man ²[the gardener ³[the boy saw ³]hit ²]is in the house.]¹

The latter example of a self-embedded construction contains two layers of centre-embeddings, which both interrupt the structure of the superordinate clause. Sentences with only one layer of centre embedding are quite common in English, as is illustrated by the following example:

¹[The purse ²[the woman left ²]is on the table.]¹

Sentences which contain two layers of centre-embedding are on the borderline of acceptability. However, sentences with three or more layers are manifestly unacceptable as sentences of English. Chomsky and Miller (1963) argue that multiply self-embedded constructions cause such difficulty because of the way the speaker's processing - strategy is organised. For example, they maintain that the human processing mechanism can only with great difficulty perform a particular operation if it is already in the middle of performing the same operation. This explanation is plausible only insofar as one accepts the transformational approach to language structure. Infinitely-recursive self-embedded constructions, for instance, are perfectly grammatical as the output of a formal grammar, as are infinitely-recursive left- and right-branching structures. But the putative correspondence between the output of such formal language grammars and natural language is an as yet unproven one. Moreover, if the only reliable

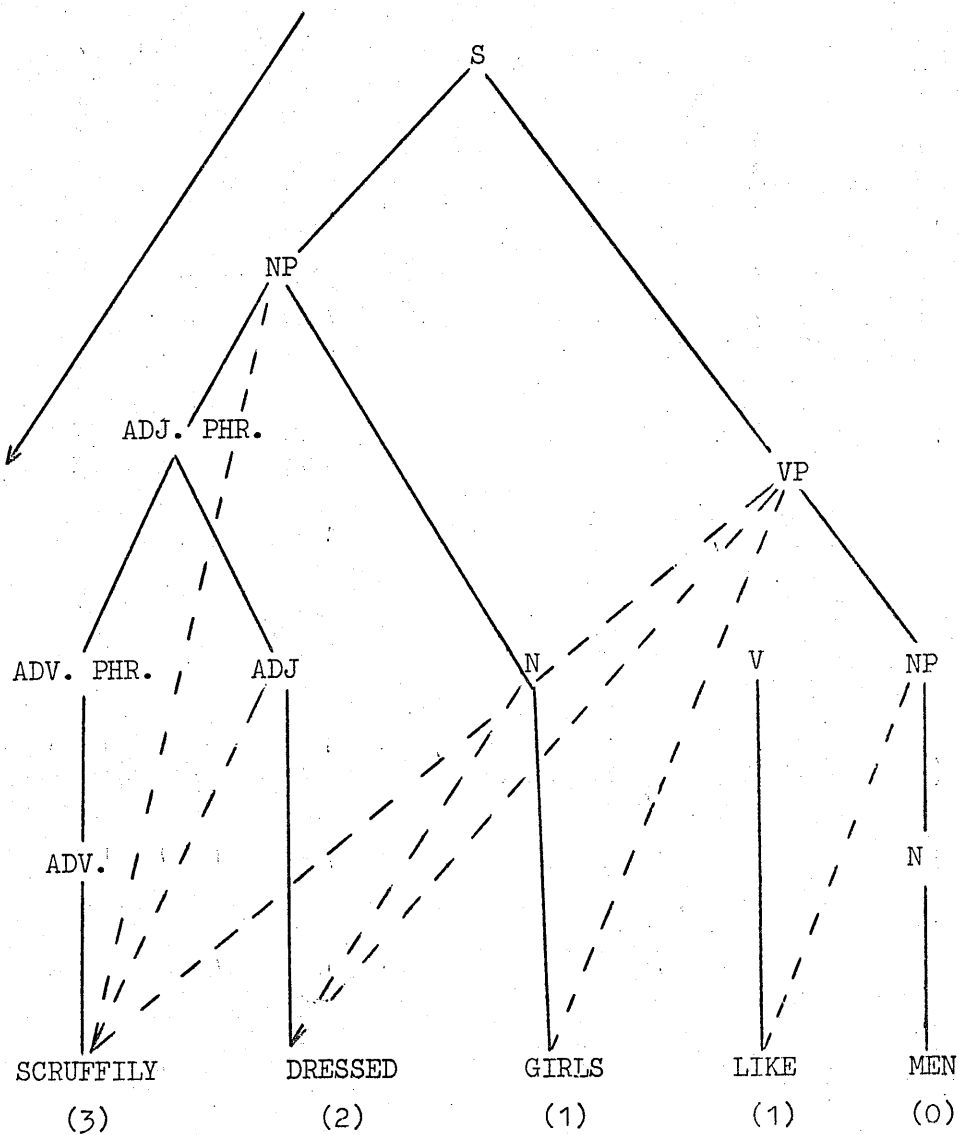
data we have for grammatical description is the speaker's linguistic performance (see, for example, Labov, 1975), then the outputs permitted by formal grammars are irrelevant as criteria for the acceptability of sentences. For these reasons, it is suggested that there are strict limits to the degree of branching permitted in English sentences. Right-branching structures, however, are preferred to left-branching structures (such as the long subject) in cases where the depth preceding the main verb exceeds an acceptable limit. The difficulty caused by excessively-long subjects is evidence of this.

It is useful at this point in our discussion to compare the depth hypothesis of Victor Yngve (1960, 1961). Yngve, too, was concerned with the structural depth permitted in English sentences. The depth hypothesis distinguished between left- and right-branching structures in sentences and employed the traditional I.C. approach to the classification of the sentence (see Introduction pp. 3-7) Yngve's aim was to produce a computer programme to generate sentences for machine translation. Because of its preoccupation with the artificial generation of sentences, the depth hypothesis is particularly concerned with what a computer has to store in memory to carry out a programme. Like the formal grammars we have just examined, the rules of Yngve's programme expand auxiliary symbols into terminal symbols (words). The programme rewrites the initial symbol and follows this with the next symbol. The programme then progressively follows each left-hand branch until it arrives at the first word of the sentence. After generating the first word, it then retraces its path up the tree expanding each right-most symbol; consequently, it must store in the short-term memory any unexpanded symbols for future rewriting. The number of items which have to be stored in this way determines the structural depth of each word as it is generated.

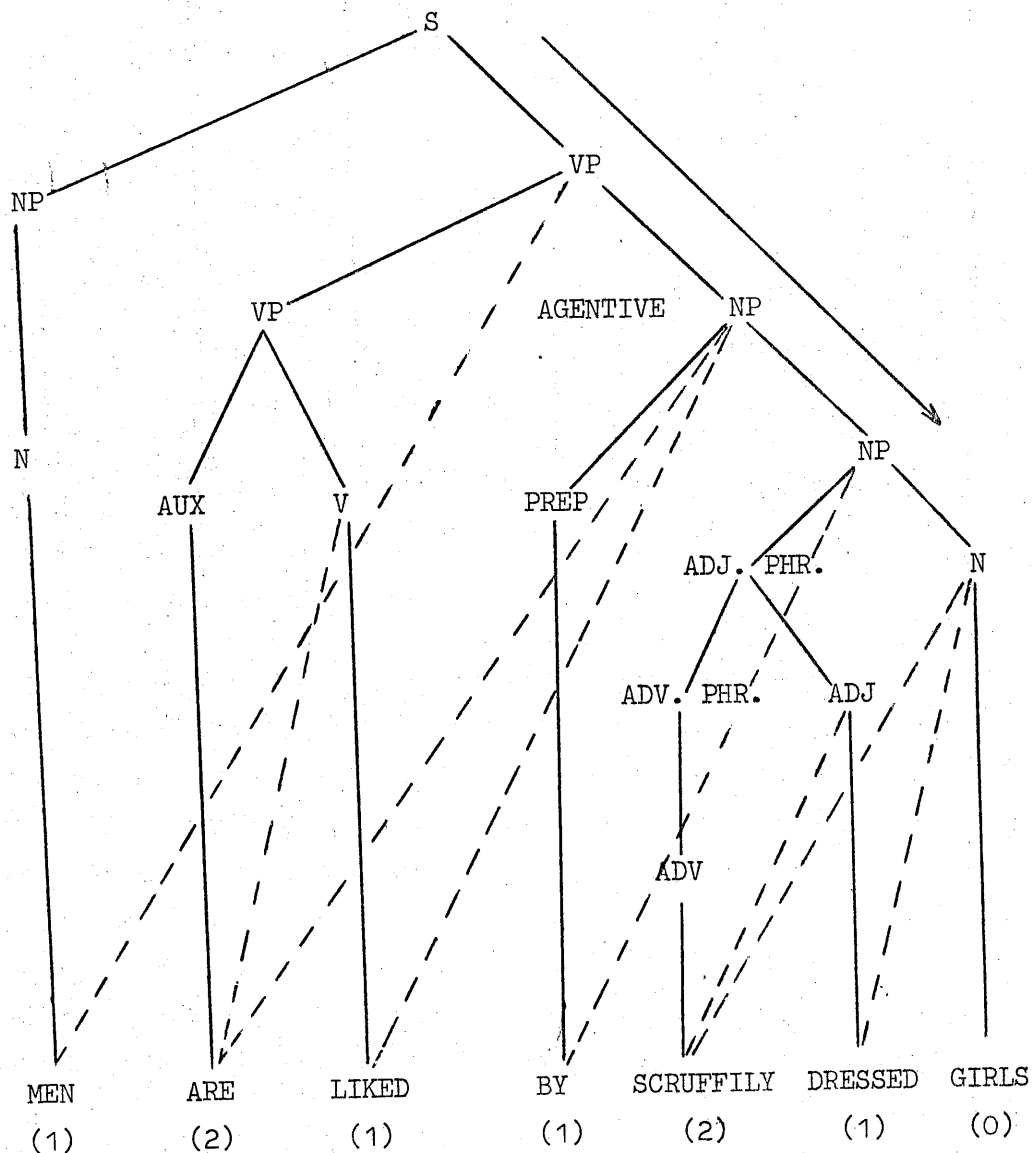
Yngve called left-hand branching structures regressive and right-hand branching structures progressive. As an example of a regressive structure in terms of Yngve's model, consider the tree on page 41, (example 1). The diagonal arrow indicates the direction of expansion in the tree. The structural depth for each word is given in brackets underneath, while the dotted lines indicate the remaining non-terminal symbols still to be expanded for each word. Yngve argued that with a finite memory, there must be some limit to the number of these symbols that can be stored at any one time. In the case of the human-language-user, he maintained that the structural depth of a word is limited to about seven items. Moreover, the

type of constructions that lead to greater or lesser depth depend on whether the sentences are right-branching or left-branching. In the case of a regressive structure, more items have to be kept in short-term memory than is the case in a progressive structure. For example, the right-branching (passive) equivalent of the structure just given would be as in (Example 2, P.42).

EXAMPLE 1



EXAMPLE 2



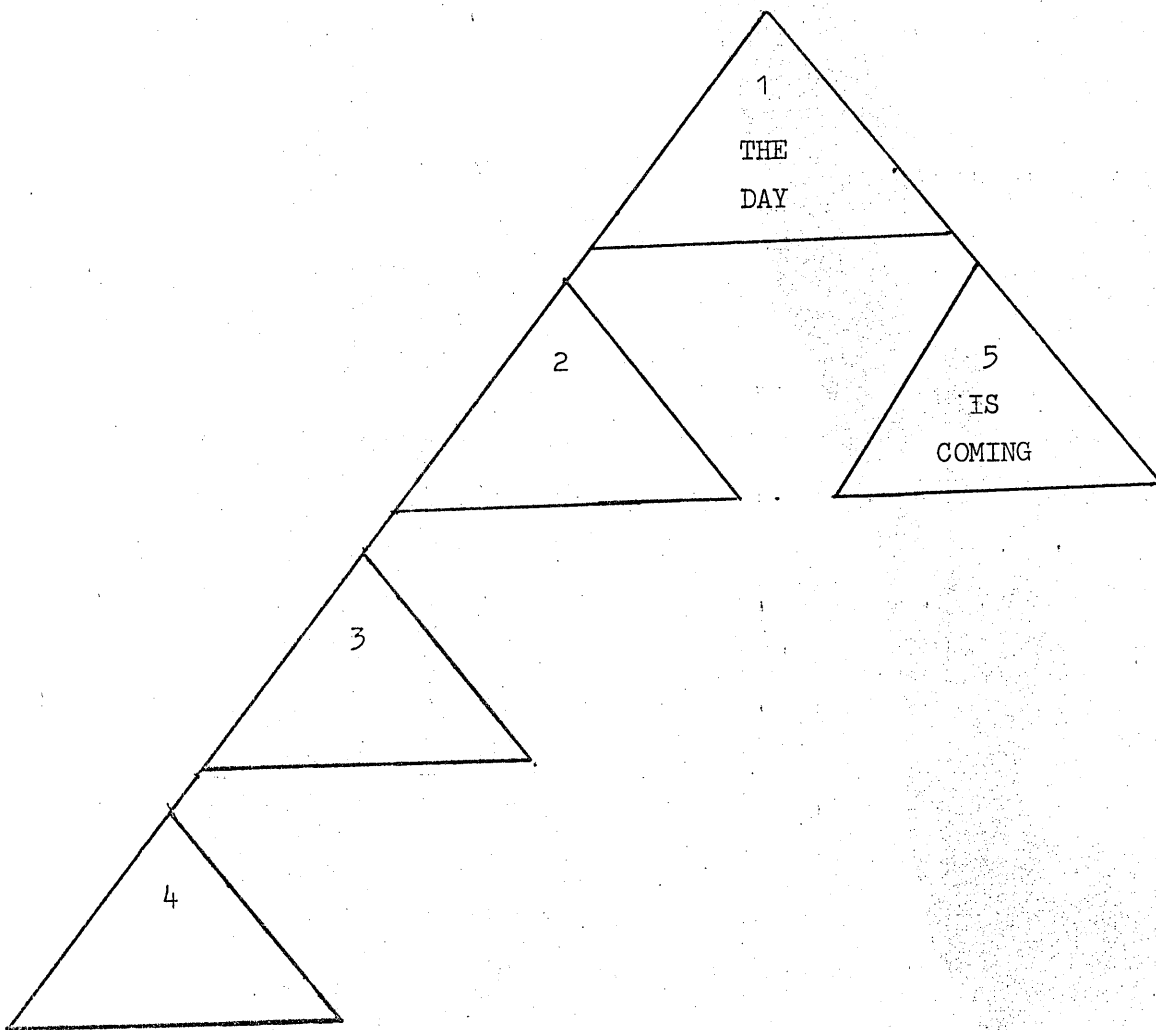
As the tree shows, this right-branching equivalent has a lower structural depth than the first (active) sentence. On this basis, Yngve claimed that the syntax of English constrains the amount of depth for each word as it is generated and provides for certain stylistic options which reduce this depth. Thus one such stylistic option would be the passive construction just illustrated. Another would be the various constructions which function to postpone excessively long left-branching structures, what we have called interruption of the long subject in section 1.1. Unfortunately, this model of speech production has certain weaknesses. It is by no means certain, for

example, that speakers proceed in sentence production in the way described by Yngve. When producing the first word of a sentence they may, in certain cases (either through a knowledge of the language or through linguistic dependencies of some kind), know how the sentence will unfold.⁴ It is not clear that speakers produce sentences one word at a time. When the sentence is considered from the decoder's point of view, there are various suprasegmental properties of language which may provide cues to the forthcoming structure of a sentence.⁵ This is in fact suggested by Martin and Roberts (1966).

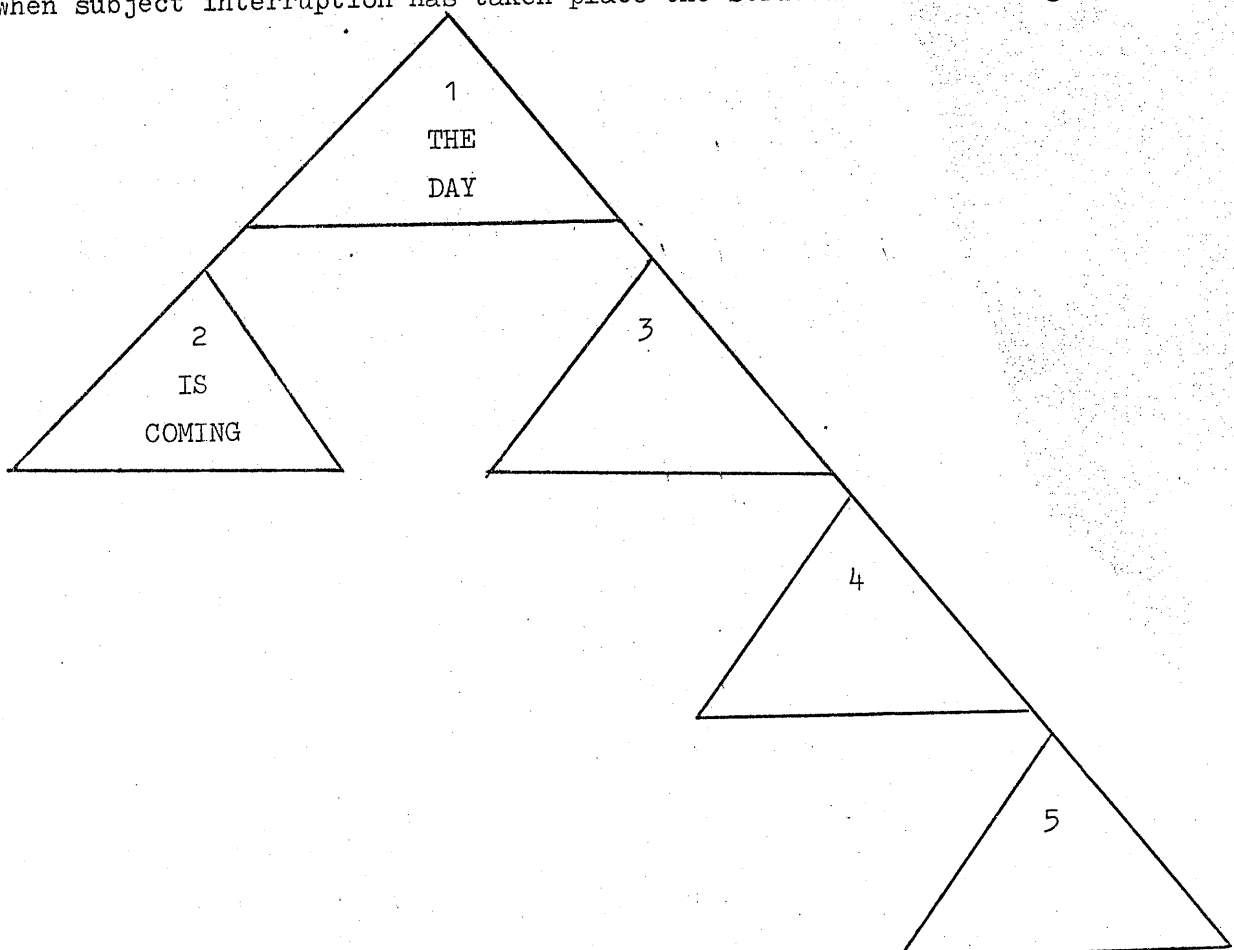
What is interesting about Yngve's depth hypothesis from our point of view, however, is the importance it attaches to left- and right-branching structures in sentences.⁶ In the first section of this chapter, we considered the use of the long subject construction (one type of left-branching structure) and the constraints governing its excessive use. In the context of this discussion, depth was related to the level of subordination preceding the main verb. Moreover, this notion of depth is interpreted much less rigorously than it is in Yngve's model. There can, for example, be no statistical maximum of structural depth in the long subject construction. Nevertheless, we are guided in the use of the long subject by an intuitive rule of acceptability to the reader (or listener). In support of Yngve, it is suggested that there do exist certain linguistic options in English whose function is to reduce depth at the left-most part of the sentence, even though this may involve a concomitant reduction of emphasis in the subject. Such a device is provided by subject interruption of long clausal subjects in sentences with intransitive verbs.

The criteria chosen for left-branching structures in this study is the expansion of structure which occurs before the main verb of the sentence. In illustration of this, we may reanalyse example (6) above (see section 1.1) according to the type of branching it employs, as follows:

5	4	3	2	1	
[[[[[The day]
					when it will not be worth making
					a statement about style in painting or
					literature]
					unless the statement can be
					so precisely formulated]
					that it can be
					translated into a computer programme]
					is coming.]



When subject interruption has taken place the structure is rearranged to:



This constraint of subject interruption for excessively long subjects is acknowledged by Quirk et. al. (1972). Quirk suggests that "it is a dominant tendency of syntactic structure that the greatest depth of subordination is reached in the final part of the sentence" (p. 792). This dominant tendency of English for right-branching structures of subordination is simply a function of the need for clarity of syntax.

In sum, this section has examined the notion of sentence depth in terms of the expansion of structure at some point in the sentence: the left-most part, the centre or the right-most part. For most practical purposes, the kind of branching which occurs can be determined from its relation to the main verb of the sentence. Within the context of a formal grammar in T.G., describing linguistic competence, there is no limit to the degree of branching which may occur in a sentence. However, from the point of view of a performance-based grammar, there are strict limits to the extent of the branching that is allowed. Right-branching structures are preferred beyond a certain limit. A reason for this preference is suggested by the depth hypothesis of Yngve, although not all of the assumptions made by the latter can be unreservedly accepted.

SECTION 1.3

The interruptions of the subject that we have considered so far (see section 1.1) are open to sentences whose main verbs are intransitive. There is, however, another way of delaying the subject and this applies to passive clauses, participle clauses and S V C clauses in which V is the auxiliary verb be and C is an adjective. According to this method, we simply invert the word-order of the clause. It is worth noting here, as noted by Huggins (1977), that cases of subject interruption from an active clause to a passive clauses may result, in some cases, in a less 'difficult' sentence. One reason for passivisation is to avoid having a subject nominal group separated from its main verb by long modifying structures, such as relative clauses. By placing these modifying structures at the end of the sentence (as in a passive construction), the memory load on the reader is reduced. This would seem to be strong evidence against the theory of derivational complexity.

Consider the following two passive constructions:

- (1) A new research laboratory built by Hitachi at Mito for testing lifts and for reporting on their behaviour is illustrated below.
- (2) Seven of the clothes that represent the best of Paris in London, from the expensive to the cheap, are drawn here.

If we represent the passive clause by S Vi Vii, where Vi the auxiliary verb be and Vii is the past participle, then the inversion of the long subject in the passive is Vii Vi S. When we apply this rule to examples (1) and (2), the following constructions are produced:

- | | | | | | |
|-----|----------------------|---|--|----|---|
| | Vii | | A | Vi | S |
| (3) | Illustrated below is | [| a new research laboratory built by Hitachi at Mito for testing lifts and for reporting on their behaviour. |] | |
-
- | | | | | | |
|-----|----------------|---|---|---|--|
| | Vii | A | Vi | S | |
| (4) | Drawn here are | [| seven of the clothes that represent the best of Paris in London, from the expensive to the cheap. |] | |

Whenever a passive is inverted, it requires an adjunct of some kind (e.g., here, below etc.) to be interposed between Vii and Vi for the clause to be acceptable English. In examples (3) and (4) the adjuncts are denoted by

the superscript A, while the long subject is shown in square brackets. These examples illustrate how the inversion of the clause in this way is a function of the long subject in English; that is, like subject interruption, clarity of syntax requires that we delay excessive depth in passive structures till after the main verb. This fact is illustrated if we shorten the subject, as in examples (5) and (6) below:

Vii A Vi
(5) Illustrated below is [a new research laboratory.]

Vii A Vi
(6) Drawn here are [seven of the clothes.]

By their strangeness, examples (5) and (6) illustrate the interdependence between inversion in passive structures and the need for clarity of syntax. Many adjectives, too, behave in a similar way to the past participles illustrated and drawn; they also take adverbial modification of some kind when the normal word-order is reversed. This is illustrated by the following incomplete construction:

V
Even more significant are ... (SUBJECT)

A more intelligible illustration is provided by example (8) in section 1.2. Using the method just introduced, this sentence can be rewritten in the following way:

(7) More convincing, however, is the explanation advanced by Phillip J. Klass in an article in Aviation Week (Vol. 85, No. 8), in which corona discharges, in the shape of balls of gas thrown up by the power lines, are held responsible for the UFOs themselves.
(original version of text)

It is clear that if we were to omit the adverbial premodifier from the example (7) (i.e., if we delete the intensifier more from the beginning of the sentence), this would produce an odd-sounding example.

There exists another (though less usual) method of improving the clarity of excessively-long subjects in passive structures. The subjects of sentences (1) and (2), for example, could be interrupted in the way discussed in section 1.1. The syntax of example (1) could be changed by inserting the sequence Vi Vii A between the nominal group a new approach laboratory and the verb built. Similarly, in example (2) we could insert the sequence Vi Vii A between the nominal group seven of the clothes and

Long subjects like the ones illustrated in these examples can be clarified by, firstly, fronting Anticipatory*- It (s_1) and, secondly, shifting the long subject to a position after the verb and its adjunct (or complement).⁷ In examples (13) to (15) below, the delayed subject is denoted by S_2 :

- (13) S_1 V C [S_2 It is safe to say that receptors function by reporting external conditions in terms of a frequency code of nerve impulses.]
- (14) S_1 V C A [S_2 It was obvious from the start that you would have taken such a negative attitude towards our proposal.]
- (15) S_1 V C [S_2 It is doubtful whether they should have analysed the problem with their limited resources and lack of adequate concepts.]

The additional superscripts denote verb and complement. Once again, by shifting the subject in this way, there may be a consequent reduction in the markedness of the subject.

In this section, therefore, we have examined two types of change that must be made to excessively long subjects. Firstly, we have seen that in passive clauses we invert the clause pattern S Vi Vii to Vii Vi S. Secondly, the long subject may be shifted and the total sentence structured by Anticipatory - It, where S_1 is anticipatory - It and S_2 the delayed subject structure which follows the complement. There seems to be some support for our argument of the relative difficulty of sentences whose main verb is delayed from the work of Limber (1973). Using children as subjects, Limber's experiment was concerned with the use of relative clauses. It was found that the children regularly used relative clauses which they attached to objects of main clauses. However, the children never used relative clauses to interrupt the main verb of the (main) clause.

* The term Anticipatory is taken from Quirk et al (72), where it is discussed in Section 14.23 as anticipatory subject. What is said here of Anticipatory It applies in principle to the signalling function of Existential There, except that instead of signalling a clause of some kind to follow its main verb, it signals a nominal group of some kind. Consider the rewriting of example (10) on Page 34 above, which illustrates this signalling:

- (10 A) S_1 V [S_2 There are courses which do an all-round job and which provide the student with powerful motivation for learning.]

Summary

This chapter has considered one use in English of marked syntactic structure, namely, the long subject construction. Although this is a very useful means of achieving emphasis in the sentence, its excessive use is controlled by a superordinate need for clarity of syntax. The latter requires that excessively long subjects be interrupted or delayed so as to reduce the perceptual load on the reader. Accordingly, the main verb of the clause is presented as the primary constituent as far as comprehension is concerned. In addition, it is possible to determine what type of branching occurs in the sentence from its relation to the main verb. English has a tendency for right-branching structures when the main verb is delayed by too much subordination. The latter is offered as support for Yngve's assertion that English has in-built alternatives to left-branching structures. Such performance-constraints on the degree of branching permitted in different parts of the clause, however, are taken as evidence against any assumed relation between the recursive branching allowed by formal grammars and that which is possible in English syntax.

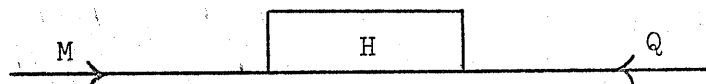
Chapter Two: The Structure of Nominal Groups

SECTION 2.1

In chapter one we examined the structure and function of the long subject construction in English text. It was necessary to discuss the role of this construction in relation to two important aspects of textual organisation, namely clarity of syntax and emphasis. While emphasis was seen to be frequently successful as a device for marking certain parts of the sentence as semantically important, clarity of syntax was seen to be a much more basic requirement from the reader's point of view. It was also shown how the long subject can itself contain other structures which further specify the meaning of the noun head. Although such structures were not examined in any great detail, it was clear that they either premodify or postmodify the noun head. For convenience, the noun head and the premodifying and postmodifying structures which accompany them will be collectively referred to as nominal groups. While one source of difficulty in text may derive from excessive use of the long subject, yet another may derive from complexity within the structure of the nominal groups themselves.¹ In what follows, it will be argued that a reader may well tolerate fairly considerable modification before or after the noun head but will experience great difficulty with excessive modification which occurs in both positions at once. This section will therefore be concerned with clarity of syntax from the point of view of complexity within the nominal group. Finally, an attempt will be made to list those structures which typically occur before and after the noun head and to show how, when misused, they may adversely affect the clarity of the sentence as a whole.

In chapter one the main terms introduced were: Subject (S), Object (O), Complement (C) and Adjunct (A). (This ignores the ubiquitous prepositional phrase as 'carrier' of the nominal group). Such terms, however, are only general labels for structures which are often more complex than might appear at first. Thus, it is possible for any of these structures to contain a nominal group. In the analysis of the nominal group in what follows, we will denote the premodifying items of the noun head by M, the noun head itself by H, and the postmodifying clause(s) by Q (qualifier). A nominal

group will therefore broadly consist of the following type of structure:



The arrows illustrate the direction in which the various modifying structures are pointing; that is, in the direction of the noun head contained within the square box. Further, we may say that the modifying items at M predict the occurrence of the noun head at H. (It is, of course, possible to have nominal groups which have zero realisation of M and Q.) Clearly, the head in the nominal group is the central (and hence, most important) semantic item. In fact, when we read a 'difficult' sentence it is to the heads that we refer for our basic semantic interpretation. In this sense, the word 'head' should be construed in quite broad terms. It is not only nouns which can be used to form grammatical groups; adjectives and verbs, for example, can form adjectival and verbal groups, respectively. The heads of these groups are referred to as lexical heads. In the case of the nominal group, consider the heads in the following sentence

H Q
H
 It is the poor (who are to be pitied) not the rich Ø

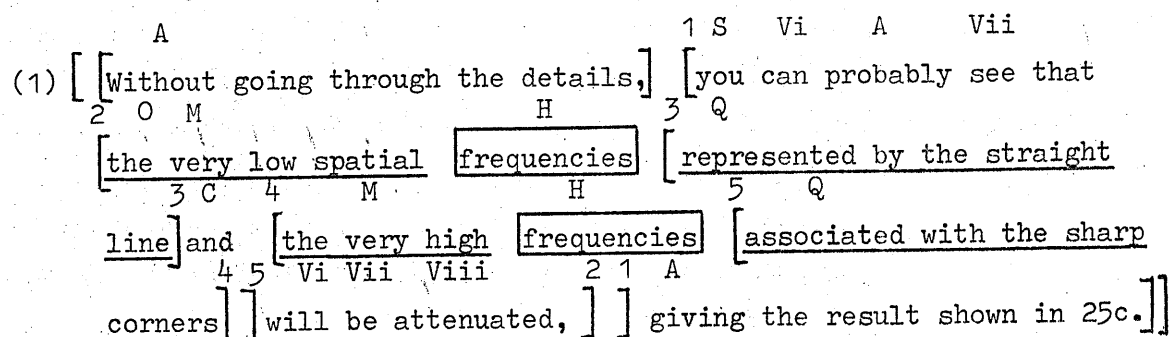
In this example, the definite article and the adjective heads are underlined and the postmodification (the 'cleft' or marked relative clause) is shown in brackets. The zero-symbol Ø denotes the deletion of the relative clause 'who are to be pitied' from the second head rich. The adjectives rich and poor are heads because they have the definite article the as premodification. In other words, the two adjectives are here functioning as nouns. The gerund or verbal noun can also be a head in the nominal group:

M H Q
 The teacher's hitting (of the boy) shocked the whole school.

More technically, such groups are referred to as endocentric constructions; that is, constructions where the head performs the same syntactic function as the whole group and may stand in place of it (Hockett, 1958, pp.183-190).

So far, it has been argued that the main semantic value of the sentence is carried by the heads. Going even further, the meaning of a sentence may be characterised as consisting essentially of the relationship between 'headed' structures. In the case of the nominal group, the purpose of the various modifying structures is to amplify and further specify the meaning of the

heads. Accordingly, it is often possible to delete this modification (i.e., either M or Q, or even both at the same time) from the nominal groups and yet still retain a 'grammatical' sentence. This is not to suggest, however, that a sentence so changed will be synonymous with the original, only that the broad semantic value will be retained. In illustration of this, we may consider the following sentence (whose constituency structure is shown in square brackets):



(ST. 291. 1 & 2)

In this example, which displays left-branching structure, the superscripts indicate the adjunct, subject, verb and object. This is the level of analysis used in chapter one. However, when we refine the analysis we see that the (long) object contains two co-ordinated nominal groups. The two nominal groups and verb phrase (i.e., Vi Vii Viii) together constitute the object of the verb 'see' in the main clause. The heads of the nominal groups, together with the accompanying modification at M and Q, are indicated by the notation introduced on page 52. We may delete the postmodification (Q) from the nominal groups and produce the following minimal grammatical form:

- (2) Without going through the details, you can probably see that the very low spatial frequencies and the very high frequencies will be attenuated, giving the result shown in 25c.

Although this new sentence does not contain anything like the specificity of the unreduced sentence, it nevertheless retains what the original was 'about': a certain state (attenuated) is predicated of the subjects (frequencies) and, as such, produces the minimum form of structure which the reader can still accept.

(3) the very low spatial frequencies \xrightarrow{H} \boxed{H}

(4) the very high frequencies \xrightarrow{M} \boxed{H}

(5) very low spatial. the frequencies.

(6) high the very frequencies.

D	O	E	N	N
determiner or predeterminer	ordinal or cardinal numeral	term of evaluation (adjective)	noun	noun head

(7) (All the^D)(ten^O)(fine old^E)^N stone houses. (A.A. Hill Pages 230/239)
VI V IV III II 1 (Hill's sequence numbering)

(8) Fine old stone all the houses.

In a Chomskyan grammar, position E allows for recursion; that is, it has no restriction on the number of evaluative items which it may contain. However, once again, when linguistic performance is taken into account, there are obvious limits to the number of adjectives which may appear in premodifying position. The number of adjectives which may appear at position E depends on the number of postmodifying clauses which follow the head (position Q). We would have great difficulty in reading a sentence

which contained a noun head preceded by, say, six adjectives and followed by five postmodifying clauses. As with the subject of the sentence, the numbers of premodifying items which may appear at position E is limited by an intuitive rule of acceptability to the reader: too much modification at M and Q will obscure the noun head. This in turn affects the clarity of the whole sentence.

If we recall example (1), it can be seen that the two postmodifying clauses of the subjects frequencies themselves each contain a noun head, line and corners, respectively. This is illustrated in examples (10) and (11) below:

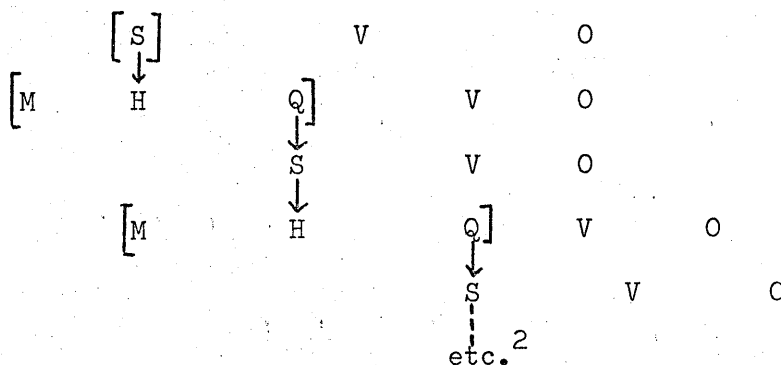
(10)

H	Q	M	H
frequencies	represented by the straight	line	

(11)

H	Q	M	H
frequencies	associated with the sharp	corners	

The two heads line and corners take the attributive adjectives straight and sharp. The two qualifying clauses thus illustrate another position in the structure of the nominal group which admits multiple repetition of grammatical structure. We could add another postmodifying clause to the each of the noun heads line and corners. To do so, however, would considerably complicate the structure of the nominal group. In the same way that there is a limit to the number of evaluative items which may precede the head of the nominal group, there is a limit to the number of postmodifying clauses which can be added to the head. In this way, there is a self-limiting relation between the premodification at M and the postmodification at Q. The latter may, in fact, be postulated as an in-built performance-constraint of English which regulates the amount of modification which occurs before and after the noun head. One of the things we look for when we try to characterise 'difficult' writing is writing which allows too much lexical detail at positions M and Q in the structure of the nominal group. This idea is made explicit by the following diagram:



So far, we have delineated the basic structure of the nominal group in terms of M, H and Q. Much of this discussion has been spent describing the principal characteristics of M and H. In the case of the former, we have noted the sequential character of its typical constituents; in the latter, we have given the criteria for 'headedness' in the nominal group. Three kinds of lexical item have been seen to function as heads: nouns, adjectives and gerunds. Until now, however, we have only glossed the structures that typically occur at Q. In the remaining part of the discussion, we will be concerned with the classification of such postmodifying structures. Finally, we will apply the notion of the nominal group presented here to an example from poetry. It is hoped that the latter will illustrate even more clearly the potential of the nominal group for complex syntactic structure.

Probably the most common type of qualifier of the noun head* is the prepositional phrase:

- (12) M H Q
The judge's attitude towards the accused was totally uncompromising.
- (13) M H Q
His involvement in the illegal trafficking of drugs promises to pave the way for his downfall.

A more diverse group of structures which qualify the head is that of non-finite clauses. For example, when the head of the nominal group is an abstract noun, the structure of the postmodifying clause is sometimes that of preposition plus gerundial clause:

- (14) M H Q
His role in rallying the support of the electors was recognised by even his opponents.
- (15) M H Q
His preference for arranging things with the greatest economy of effort often contributed to his success as an administrator.

Other non-finite constructions which follow the head are those which contain a present participle:

- (16) M H Q
The detective is the man standing in the vestibule.

* A point long recognised in studies of nominal heads is that abstract nouns like those in (12)(13)(14) and (15) are potentially more difficult to comprehend than concrete nouns like man in (16).

or a to-infinitive:

M H Q

- (17) The problem to address yourselves to has already been mentioned.

or finally, a past participle:

M H Q

- (18) The applicant rejected by the interview panel decided to look elsewhere for employment.

Adjectives may also follow the head:

M H Q

- (19) The student most anxious to work is not necessarily the most intelligent.

The latter is, strictly speaking, an adjective plus a to-infinitive.

A variety of adverb forms may also follow the head:

M H Q

- (20) Conrad's experiences in Africa are at the heart of his novels.

M H Q

- (21) We spent our holidays last summer on the Norfolk Broads

Relative clauses are another common way of qualifying the head (formed by who, whom, which, that, whose, etc):

M H Q

- (22) Their handling of the problem which I had found quite impossible was a success.

There are also constructions which contain nouns in apposition to the subject of the sentence. Where the item in apposition is a clause, it is called an appositive clause (Leech, 1975, p. 270). In the case of the simple noun head in the postmodifying clause, however, we say that there is modification by a second noun head:

H M H Q

- (23) 'Lee Harvey Oswald, the assassin of President Kennedy, had access to highest priority secrets when he was a marine....'

(Guardian, 20.2.78)

The postmodification in example (23) contains an endocentric head and the apposition relation is therefore between two endocentric structures. This

is to be contrasted with example (24), which has a noun head followed by a subordinate that clause:

- M: H Q
(24) The assumption that differences between different populations or different ecotypes would disappear if they were subject to a common environment is a hangover from the early descriptive phase of ecological genetics.

(ST. 291 (1 & 2))

In example (24) the appositive clause, together with the subject (the assumption) creates a nominalisation. This nominalisation takes the form of turning a finite clause into an abstract nominal group. For example, the corresponding finite clause to example (24) would begin:

People assume that

Examples (12) to (24) represent the main types of postmodification of the noun head. To complete this discussion, we may now apply our total characterisation of the nominal group to an example from poetry. Consider, for example, the following extract from Wordsworth's Tintern Abbey:

And I have felt
A presence that disturbs me with the joy
Of elevated thoughts; a sense sublime
Of something far more deeply interfused,
Whose dwelling is the light of setting suns,
And the round ocean and the living air,
And the blue sky, and in the mind of man;
A motion and a spirit that impels
All thinking things, all objects of all thought,
And rolls through all things.

The first thing we notice about this extract is that it provides an example of what has been called the verse paragraph (Widdowson, 1973). That is, the extract constitutes one complete sentence which spreads over many lines (10 in all). The second thing we notice is that the total sentence consists, broadly speaking, of a subject (I) and a transitive verb (felt) which takes three nominal groups as direct objects, each of which is quite complex in structure. The final (object) nominal group, for example, embeds two further object nominal groups ('All thinking things, all objects of all thought'). These latter nominal groups are the direct objects of the verb impels. Thus, the structure and complexity of

this extract can be explained in terms of the structure of its constituent nominal groups. The basic structure of the extract is given in example (25) below:

- S Vi Vii
- (25) 1. And I have felt
- M H Q M H
2. Oi(A presence that distracts me with the joy
- Q H Q H M
3. Of elevated thoughts;)Oii (a sense sublime
- Qi H Q
4. Of something far more deeply interfused,
- Qii
5. Whose dwelling is the light of setting suns,
6. And the round ocean and the living air,
7. And the blue sky, and in the mind of man;)
- M H H Qi
8. Oiii(A motion and a spirit that impels
- M H H Qi
9. (All thinking things, (all objects of all thought,
- Qii
10. And rolls through all things.)))

Embedded
objects

In example (25) the Q superscripts denote the postmodifying clauses, and the other superscripts (i.e. M, H and Q) denote the embedded modification within the clause. Clearly, not all of the smaller nominal groups are shown in this analysis; for clarity of presentation, only the superordinate nominal group structures are indicated. The square brackets indicate the total length of the (inclusive) postmodifying clause, while the round brackets show the length of each of the object nominal groups. As can be seen, most of the postmodifying clauses at Q are relative clauses (see example (22) p. 57). The relative clauses begin on lines (2), (5) and (8) with that, whose, and that, respectively. Multiple postmodification is illustrated in lines (2), (3), (4), (5), (7) and (9), with the embedded heads and their accompanying modifiers. Frequent use is made in this extract of the co-ordinating conjunction and to conjoin headed structures (signalled by double lines). We might also note, finally,

the inversion of the normal order of premodifier + noun which occurs at the end of line (3). Here the normally attributive adjective sublime occurs in position Q, where it becomes a postposed adjective.

In conclusion, the purpose of this section has been to increase the specificity of our categories for analysing syntactic complexity in English text. This complexity cannot be satisfactorily analysed with only such broad constituents of the clause as S, O, C and A. These major constituents often contain an inner complexity of their own. Any analysis of syntactic complexity must try to characterise this internal structure. This structure has been seen to consist of a lexical (or grammatical) head with its accompanying modifiers. It is the scope for indefinite repetition within the various modifying structures that often creates syntactic complexity in the sentence. Once more, this scope for indefinite repetition is controlled by an intuitive rule of acceptability to the reader. Excessive modification at positions M and Q in the nominal group(s) will obscure the semantic relation between the various heads; a concomitant effect of this will be to delay the sentence-boundary. Perhaps the most important implication here is that large-scale syntactic complexity in text is often a consequence of complexity at a much more micro level of structure.

SECTION 2.2: Apposition

The previous section illustrated the focal importance of lexical heads in sentence structure. Next to the main verb, the heads are the most important constituents of the clause and represent its basic semantic content. The function of the various modifying structures is to make more precise or further specify the meaning of the heads which they accompany. From a strict semantic point of view, therefore, the modifying structures within the sentence occupy a more subordinate role than the lexical heads. The syntactic clarity of a sentence is impaired when its modifying structures obscure the relation between its lexical heads. A 'difficult' sentence may also result when the headed structures themselves, due to excessive modification, delay the main verb (and hence, also, the sentence-boundary). Up to this point, however, we have only considered the modification of lexical heads in terms of the premodifying and postmodifying structures of the individual nominal group. In this section we will consider another, equally important, way of modifying the heads of a sentence, namely, apposition between headed structures. In section 2.1 we briefly introduced apposition in terms of its postmodifying function (see examples (23) and (24) above) but did not examine it in any great detail. Although apposition may occur between the heads of a great variety of the elements of clause structures (e.g., between adjectives, verbs, complements, etc.), the examples of apposition we will consider will be apposition between nominal groups. The aim of this section is to show that apposition, if used excessively, may also contribute to syntactic complexity in text.

Apposition denotes both a semantic and grammatical relation between heads or headed structures. Items related in this way are called appositives. The grammatical relation between appositives resides in an identity of grammatical form; the semantic relation concerns the co-referential nature of the constituents. 'Identity of grammatical form' means that the second appositive must repeat in some way the grammatical form of the first. Such repetition may take the form of direct lexical repetition, deletion, substitution or a combination of these forms. The technical name given to such repetition of clause structure is replacement (Winter (80) forthcoming).³ 'The co-referential nature of the constituents' means that the two appositives must either be identical in reference or the reference of the one must be included in that of the other (Quirk et al., 1972). Thus,

in example (1)

- (1) Mrs Jones, the farmer's wife, was elected chairperson of the committee last week.

'Mrs Jones' and 'the farmer's wife' refer to the same person. The two are both nominal groups and so produce a relation of nominal apposition. There are two types of nominal apposition, however, which must be distinguished. The first, as illustrated by example (1) above, is where the second head of the relation does not repeat the lexical item of the first head. In the second type, the second head repeats the lexical item of the first head, as illustrated in example (2) below:

- (2) The President spoke of the problem, America's problem, of how to avoid engaging in a nuclear arms-race with Russia.

In example (2), the first head (problem) is lexically repeated by the second head. The replacement also takes the form of substitution of the premodifier America's for the definite reference of the first head. This relation is sometimes made fully explicit by such adverbials as more precisely, more exactly, etc.

One good test for apposition is to try and rewrite the two appositives in the form of a S V C structure. The appositives in examples (1) and (2) would therefore be rewritten as (3) and (4):

- (3) S V C
Mrs Jones is [the farmer's wife.]

- (4) S V C
The problem is [America's problem.]

In this respect, apposition is an intensive relationship (Quirk et al., 1972). If we take the first appositive as our point of reference, then its relation with the second appositive is clearly cataphoric. However, this cataphoric relation is not a defining feature of apposition and does not enable us to distinguish it from other types of grammatical relation. Consider, in this instance, example (5) below:

- (5) One thing is sure; the B.B.C. will be broadcasting the truth long after the present regime can make trouble for the Teheran correspondent.

(Daily Telegraph, 27.3.79).

In this example, the whole of the second independent clause is a lexical realisation (See Winter (77) PP.86/87 and Chapter Four) of the nominal group One thing in the first independent clause. One thing anticipates the lexical detail provided by the second independent clause. The latter simply adds information to the general statement of the first independent clause. There is no replacement of structure here, no common head between the two clauses. In example (5), therefore, we do not speak of apposition between the two clauses, but of cataphoric substitution. Apposition must also be distinguished from verbless adverbial adjuncts. This is illustrated in examples (6) and (7) below:

A

(6) Normally a quiet person, John nevertheless became indignant at his neighbour's suggestion.

A

(7) A close friend of the policeman, Paul was able to take great liberties with the law.

In example (6) the verbless adjunct is concessive ('Although he was a quiet person'), while in example (7) it is causal ('Since he was a close friend of the policeman'). Examples such as these are clearly distinguished from apposition when their lexical content is taken into account.

Like Q of the nominal group, the second appositive postmodifies the first appositive (or first head). As for M and Q, the function of the second appositive is to amplify or further specify the meaning of the first. The implication here is that the first appositive is the most central semantic item of the two; just as for M and Q of the nominal group, we can often delete the second appositive and yet still retain a grammatical sentence. An incidental effect of this second head is to interrupt the relation between the first head and the adjoining clause structure. In example (1), for instance, the second appositive ('the farmer's wife') interrupts the relation between the first appositive Mrs Jones (subject) and the main verb was. In a similar way to the long subject construction, the second appositive is here delaying the main verb of the sentence. There is thus scope in apposition for delaying important parts of clause structure; moreover, its over-frequent use in the sentence will result in a complex syntax. The likelihood of excessive apposition leading to a difficult sentence is all the greater when we reflect that it may occur between or within the structure of the

major constituents of the sentence, namely S, V, C, O and A.

Apposition has no single typographical signal; in fact, in some cases there is no punctuation used at all. Appositives may be linked by any of the following types of punctuation: semi-colon, comma, dash or round bracket. There are, in addition, certain adverbials which make further explicit the appositional meaning of the relation. Typical items of connection in this respect are: namely, that is, that is to say, this is that - clause, more specifically, in particular, to be more precise, more precisely, what this means is, for example, for instance, etc. In those cases where there is no such explicit indicator of the relation, the nature of the relation is implicitly signalled by its repetition semantics. That is, the identity between the heads of the relation denotes the appositional meaning; it is a characteristic of implicit apposition, however, that it can be made explicit by the insertion of adverbials.

In what follows, we will examine a variety of types of apposition, paying close attention to the nature of the interruption in the structure of the clause which results from its use. It is hoped that the examples analysed will be sufficiently comprehensive to allow certain generalisations to be made about the range and structure of apposition. These examples will, at the same time, enable us to characterise certain examples of 'difficult' text at the end of the section. The bulk of the examples used in this analysis of apposition are culled from Open University course units. With the exception of certain examples considered at the end of the section, however, this selection is purely one of convenience. The source of each example, together with the page reference, is given in brackets underneath. The constituency-structure of each sentence is indicated by a labelled bracketing, while the first appositive is denoted by a single horizontal line and the second appositive with two lines. Finally, any deletion which occurs in the second member of the relation is shown by the zero-symbol \emptyset in that part of structure where it occurs.

The first set of examples we will consider represent the most common type of apposition, namely, where the second appositive interrupts the relation between the subject and main verb of the clause. A variety of types of apposition may perform this same interruptive function. Consider, in this instance, example (8) on page 65.

- (8) [[^S ^M ^H ^Q
A systematic comparison of the two descriptions] (that is,
^M ^H ^H ^Q ^{Vi} ^{Vii}
some kind of content analysis of the two Ø) might have
^{Viii} ^O
given [a different impression.]]

(D. 305. Block 8, 14-15, p.52)

In this example the second appositive is typographically signalled by round brackets and is made fully explicit by the adverbial that is. The appositional meaning of this relation is one of narrowing-down; the second head provides a more precise specification of the first appositive. What is especially interesting about this example is the replacement which occurs in the second member of this relation. The latter replaces the subject of the sentence; more precisely, it replaces the subject's lexical head and modification. The head of the first nominal group (comparison) is shown in the box; so, too, is the compound noun head in the second nominal group (content analysis). The new head of the second nominal group is postmodified by the same prepositional phrase as the first, namely, 'of the two', except that the head (descriptions) has been deleted. In short, example (8) illustrates an apposition relation between two nominal groups, the second of which delays the verb phrase of the main clause. The replacement meaning of the second member is thus manifested by a combination of substitution and deletion.

Example (9) below illustrates a slightly different type of replacement of the postmodification of a nominal group in the second appositive:

- (9) [[^SMr. Pym's ^Mvision ^Hof the EEC's future^Q]] - ["] ^M ["]
^H ^Q development leading to an exciting new future at the end of the
^V ^C
road - is [as enticing as it is vague.]]
 (Guardian, 21.3.79)

The head of the subject in this sentence is postmodified by the possessive phrase 'of the EEC's future'. The second appositive replaces this postmodification by a new nominal group. The common noun head between the two appositives is vision, although this head and its premodification have been deleted in the second, apposed nominal group. Unlike example (8), the interruption between subject and main verb in this example is typographically signalled by the dash and there is no explicit item of connection between the two apposed members. Example (9) is therefore a case of implicit apposition.

The relation in example (10) is also typographically signalled by the dash:

- (10) S M H Q
- [[The actual characterisation of male or female consciousness] -
- C M H
- what it feels like to be a woman or a man and the difference
- V
- it makes to the ways in which we are in the world - demands
- O
- [considerable subtlety.]]

(D 305. Block 9. 16-17, p.10)

Example (10) illustrates apposition between a nominal group (first appositive) and a co-ordinated WH-clause and nominal group. The WH-item in the second member functions as a noun head, so that there are, in fact, two noun heads in this second member of the relation. If we take apposition as an intensive relationship, then the second member could be seen as deleting the noun head characterisation, together with its postmodification 'of male and female consciousness', and adding a complement. Thus, the two members might be rewritten: 'The actual characterisation of male or female consciousness is ...' The main verb demands is delayed by the second appositive. Finally, example (10) also illustrates implicit apposition: the explicit indication of the relation (that is) has been deleted, although the latter could, of course, be inserted without changing the meaning of the relation.

Another type of grammatical replacement is illustrated in example (11):

- (11) S M H H
- [[This rather surprising finding] - that occupations appear
- V1 V11 C
- to be listed in rank order for certain parts of the country, but
- not for others, in a country-wide survey - should serve as [a
- A
- warning against making assumptions about how lists are drawn up,
- [without very careful examination of the whole documents,] [prior
- to any sampling procedure that may be undertaken.]]]

(D. 301, 13-15, p.28)

The replacement that occurs here is of a that-clause apposed to the nominal group 'This rather surprising finding'. The intensive

relation between the two members is shown if we rewrite them as: 'The rather surprising finding is that X'. Thus, the head of the subject of the sentence is replaced by a new noun head (occupations) in the subordinate that-clause, which delays the main verb of the main clause.

A substitute this as head in a relative clause which defines it can also function as appositive, as in (12) below:

- (12) ^A[[According to Marx,] ^S[the other aspects of social organisation
Vi Vii PP
mentioned on page 11] are founded ^S[on this class structure,] ^H[so
that ^S[capitalists] ^H- ^Hthose ^Qwho wield economic power - tend to
be [the most economically privileged to control political power
and to enjoy higher status than the labouring class.]]]

(D.301. 13-15, p.14).

Here, the subject of the first member (capitalists) is replaced by the deictic item those in the second member, which is qualified by a restrictive relative clause. Thus, the subordinate clause of the second appositive interrupts the relation between the subject and verb (to be) of a (super-ordinate) subordinate clause.

The apposition relation of example (13), too, functions to interrupt the subject of a subordinate that-clause from its main verb:

- (13) ^S[[The biographies] ^{Vi} were so constructed ^A [that ^{Vii} the individual ^M items ^H
^Qcomprising them] ^M (type of ^H background, ^H social ^M status, ^H school
^H achievement, etc.) ^{Vi} were systematically and independently varied. ^A ^{Vii}]]]

(D 305. Block 8, 14-15. p.59)

In this example, the subject nominal group is more specifically replaced in the second member by three separate nominal groups, which are typographically signalled by round brackets. The appositional meaning of the relation is implicit, but could be made explicit by inserting an adverbial such as for example or that is before the first head in the second member. Once more, the second member delays the main verb, in this case were.

In example (14) below, however, the adverbial item for example does appear explicitly, thereby marking the relation as one of instantiation:

- (14) ^A [Thus, ^S ^H ^Q emotions ^H ^C ^H ^V ^C which were subjectively rated as similar to one another] (for example, ^H ^C ^H joy and cheerfulness) were [more difficult to distinguish than ^H ^C ^H emotions ^Q rated as different] (for example, joy and disgust).]

(D 305. Block 8, 14-15, p.76).

The subject of this sentence is composed of a noun head and a restrictive relative clause. The second appositive replaces the latter with two co-ordinated nominals 'joy and cheerfulness', thereby delaying the main verb were. Example (14) is especially interesting in that it contains two instances of apposition. In the second relation, the second member also consists of two co-ordinated nominals in apposition to the subject emotions. The first head in this relation, however, is part of a complement which is functioning as a comparative clause. Although there are two instances of apposition in this sentence, there is only one case of interruption between subject and main verb, in the first relation. Example (14) therefore illustrates a case of two-fold parallel apposition, where the relation in each case is both functionally and grammatically the same.

Correlative commas are used in example (15) to signal the second appositive:

- (15) ^S ^H [Privatisation, ^{Vi} ^A ^{Vii} ^O an ugly word for a simple solution to over-manning, is already attracting [attention from go-ahead authorities.]]

(The Daily Telegraph, 27.5.79).

The second member of this relation occurs within an evaluation of the first member and is limited to 'a solution to over-manning'. As we will see in chapter three, the lexical items evaluating the first member of this relation fulfil quite a different semantic function in the sentence to that of apposition. Again, the second member of the relation delays the main verb (in this instance is) of the main clause.

In sum, examples (8) to (15) all illustrate how the same interruptive function (i.e. between subject and main verb) can be carried out by a variety of appositional forms, all operating within the structure of the nominal group.

The interruption between subject and main verb may occur within the context of a main clause or a subordinate clause. A variety of typographical devices may be used to signal the same type of interruptive relation. In the examples which follow, we will see that there are many more parts of clause structure which admit interruption by an appositionive structure. For example, it is possible to have an apposition relation which interrupts an embedded adverbial adjunct from its main clause:

- (16) ^A[[Finally,] ^A[when Triandis and Hambert came to look at the emotion categories into which the subject groups had placed the expressions,] ^S [they found ^V [that ^O [while the American group made ^A fairly 'accurate' ^M ^H ^M ^H ^Q ^V ^O identifications] (i.e. the categories that they chose matched the emotions intended by the photographic model),] [the Greek groups made a significant number of misidentifications.]]

D 305, Block 8, 14-15, p.81)

In example (16), the main structure is represented by a matrix-clause of the form 'they found that X'. The function of the matrix-clause is to embed another sentence, which it then comments on or evaluates in some way. In this sentence, the embedded sentence of the matrix-clause is represented by 'the Greek groups made a significant number of misidentifications.' The latter is delayed by an adverbial clause which is embedded within the total matrix clause. This adverbial clause, however, is itself part of an apposition relation which further delays the embedded sentence of the matrix-structure. In the second appositionive, the head identifications is replaced by the head categories and its postmodifying (restrictive) relative clause. The apposition which occurs in example (16) thus interrupts the larger matrix-clause; moreover, it further delays an already delayed clause by occurring immediately after an embedded adverbial adjunct clause.

This same type of interruption by an apposition relation is illustrated in example (17):

- (17) ^S [[The most interesting feature of Little's results] ^V was [that [on the ^A ^M ^H ^Q ^M second measure of complexity] (the number of independently deployed

^H
categories) [^C female students tended to score higher than
 males, on the third measure male students scored higher than
 females, and on the first measure the scores of the two groups
 of students did not differ significantly.]]]

(D 305, Block 8, 14-15, p.44).

In this example, the embedded structure of the matrix-clause (i.e. 'the most interesting feature of Little's results was that X') functions as the subjective complement. The latter consists of three independent clauses, which are delayed by the second appositive. Once more, an embedded adjunct clause precedes the clause embedded by the larger matrix-clause. Thus, example (17) also illustrates a two-fold delaying of clause structure. The second member of the apposition relation replaces the head measure by the head categories and its premodification, which is typographically signalled by round brackets.

Apposition can also interrupt an adjunct clause from a main clause which does not occur in a matrix-structure. This is the case in example (18) below:

^A
 (18) [[When writers like Simone de Beauvoir and Sheila Rowbotham begin
 not only to delineate the constructed nature of such dimensions,
^A ^V ^O ^H ^Q
 [but also to propose [new modes of apprehending reality]]] (for example,
^H ^Q
in notions of transforming women's consciousness), [then our view
 of sections of the world and of ourselves as stable and natural may
 be temporarily undermined.]]]

(D 305, Block 9, 16-17, p.11).

The first appositive in this example functions as the direct object of the verb propose and is part of a larger adjunct clause. The second appositive interrupts this adjunct from the main clause (beginning 'then our view of sections ...'). In this example, there are thus three separate structures which all function to delay the main clause: two sentence-initial adjuncts and an apposition relation. Finally, the head modes is replaced in the second appositive by the head notions and its modification. A parallel

type of structure is provided by example (19):

- (19) $\begin{array}{ccccccc} & A & & A & & M & & H \\ & & & & & & & \\ & H & & H & & H & & H \end{array}$ [[Certainly,] [in the choice between major perceptual classes]
(e.g. attitudes, ideas, abilities, traits), [individual
perceiver factors (as well as context) must play a larger part
than the characteristics of the perceived person.]]

(D 305, Block 8, 14-15, p.45).

In this example, there are also two adjuncts (the first of which is a sentence adjunct) before the main clause (which begins 'individual perceiver factors...'). Here, too, the main clause is delayed by three levels of structure: the two adjuncts and second appositive. The example differs from example (18), however, in that the first head (classes) is replaced in the second member of the relation by four, unmodified noun heads. But like example (18), the appositional meaning is one of instantiation, explicitly indicated by the adverbial for example.

Co-ordinated structures are also frequently interrupted by appositives:

- (20) $\begin{array}{ccccccc} & & & & & M & \\ & H & & Q & & H & \\ & & & & & & \\ & & & & & H & \\ & & & & & H & \\ & & & & & H & \end{array}$ [This is true both of [our inferences about the more static
features of others] (e.g. traits of personality) and [the^M more
dynamic features (e.g. mood, emotion).]]

(D 305, Block 8, 14-15, p.54).

In this example, the correlative co-ordinators are shown by the broken line. There are two instances of apposition in this sentence, the first of which interrupts a co-ordination of two nominal groups. In the first relation, the head features is replaced by the head traits in the second member, while in the second relation, the second member adds two unmodified noun heads. Thus, although there are two separate apposition relations in this sentence, there is only one case of interruption.

A different type of co-ordination is interrupted in example (21) below:

- (21) $\begin{array}{ccccccc} & A & & & & S_1 & Vi & C \\ & & & & & & & \\ & S_2 & & Vii & & O & & M \end{array}$ [[In discussing the inferences we draw about others,] [it is useful
[to distinguish between [those which concern relatively static

H
Q
H
H

unchanging
features
of others
(e.g. personalities, attitudes,

H
C
O
M

abilities
)
and
[these which concern relatively dynamic changing

H
H
H

features
(e.g. thoughts,
emotions
and moods
).]]]]

(D 305. Block 8, 14-15, p.55)

The main clause of this sentence begins with Introductory-It (S_1) followed by a complement and the real subject (S_2). The latter contains two co-ordinated object nominal groups, both of which are objects of the verb distinguish. Both objects are each apposed to a second appositive which consists of three unmodified (plural) noun heads. It is only the first apposition relation, however, which interrupts the structure of the clause. In this relation the second appositive interrupts the co-ordination of the two object nominal groups. In sum, this example illustrates an apposition relation which is used to interrupt the structure of a delayed subject. We might also note, finally, that the second appositive in both relations occurs at the end of a relative clause.

A final instance of the interruption of co-ordination by apposition is provided by example (22):

A
A

(22) [[However,] [as the months went by] (overall we visited the wing

for a period of four years) [we began to lose interest in the

nature of their criminality and became increasingly absorbed by

the problems which they faced in coping with their present reality;]

M
H
Q

[not so much the evident physical properties of living in the wing -

M
H
M
H
M
H
Q
M

the locks, the landings, the guards on every corner, the

H
M
H

oppressive heat, the lack of ventilation - but rather the type

of inner subjective life which developed in response to those

oppressive circumstances.]]

(D 305, Block 9, 16-17, p.13).

At first sight, this example contains only one instance of apposition: the head properties is replaced in the second appositive by five separate nominal

groups, which complete a relation of instantiation. On a closer examination, however, the whole of the second part of the sentence (i.e. that part after the semi-colon) can be seen as supplying the particulars of the 'problem' referred to in the first part of the sentence. This second part of the sentence can therefore be seen to be in an apposition relation with the head of the first part (problem). In sum, there are two instances of apposition in this sentence, where the second instance is embedded within the first. Moreover, the second relation interrupts the correlative co-ordinators not only but also, which are used to co-ordinate two nominal groups.

Another case of double apposition is provided by example (23):

(23) ^A ^S ^V ^C
^C ^A ^S ^{Vi} ^{Viii} ^M ^H
 and [in general] [a person's class] is determined by one crit^Q ^M ^H ^Q
only: his position in relation to the means of production; in
other words, whether he is an owner of capital employing labour
in the pursuit of profit, a non-owner who has to sell his labour
to a capitalist in order to survive, or a non-capitalist landowner.

(D 301, 13-15, p. 11).

The third appositive in this example is shown by the three horizontal lines. The head criterion in the first appositive is replaced by the head position in the second. The latter, however, is replaced by three separate nominal groups in the third appositive, these heads being owner, non-owner, and landowners, respectively. The explicit item of connection for the third appositive (in other words) suggests that this third relation is one of re-phrasing.

All the examples that we have considered so far have been of appositional relations within the same sentence, even cases of double apposition such as examples (22) and (23). Apposition can also occur across sentence boundaries, where the first and second appositives are in different sentences. In such cases, the appositional meaning of the relation is often made explicit by an adverbial of some kind. Thus, consider example (24):

(24) [It seems likely that [^{S_M} the ^H ^Q categories which are 'salient' to the

individual^C are [those with 'functional significance']]
 (Tajfel, 1969).] [In other words, [the salient^S categories]
^V are [those which are of most use to the individual in organising
^C his world.]]

D 305, Block 8, 14-15, p.41).

Clearly, when apposition does occur across sentence boundaries there is less deletion of structure and more repetition of the structure of the preceding appositive. For example, in the second appositive (sentence (2)) there is a direct repetition of the S V C structure of the first appositive. In addition, there is also repetition of the noun head categories and its premodification (salient). However, while in the first appositive salient is part of a postmodifying relative clause, in the repeated structure it becomes attributive to the noun head categories. The repeated complement of the second appositive rephrases the complement of the first member. Nominal groups are still the basis for such repetition; in example (24) the nominal groups represent the subject of each sentence. In such cases of apposition (which occur across sentence-boundaries) we can no longer speak of interruption of the structure of the clause. As we will see in Chapter four, this type of apposition interrupts larger discourse structures which are responsible for the relations between sentences.

There is a similar repetition of structure in example (25) below:

(25) [The judgement studies,^S] [like Darwin's studies,^A] assume^V the^M
^H ^C ^M ^H ^Q
determinacy and discrete quality of emotional states and their
associated expressions.] [In other words, in posing emotional
^S ^V ^O ^M
 states, [Feleky et. al.]^H assumed that [they were obtaining 'the
right' emotion poses, and that those used were in real life uniquely
associated with the different emotions aimed at by the models.]]

D 305, Block 8, 14-15, p.74).

In contrast to example (25), in this sentence it is the object (nominal group) of the verb assume which is repeated in the second appositive sentence.

its direct lexical repetition, example (27) is even more marked than example (26). There are two reasons for lexical repetition between adjoining sentences. The first represents an attempt to avoid ambiguity of reference. The second, however, has to do with the rhetorical marking of certain parts of structure, as illustrated in example (27).

The examples of apposition that we have analysed so far enable us to make certain generalisations about the structure and function of apposition within the sentence. Firstly, replacement semantics is an obligatory part of appositional structure. This replacement is a function of an underlying semantics of rephrasing or narrowing-down of the lexical content of some part of clause structure. Viewed in this way, apposition is a postmodifying structure whose scope of backward reference is to the nearest identical grammatical or lexical structure. As we have seen, heads constitute the principal items of appositional structure; it is around such items that replacement takes place. Secondly, a variety of typographical devices are used to perform the same appositional function, although apposition cannot be defined in terms of any one of these. Thirdly, the implications of apposition for syntactic complexity in text arise from its interruptive role within structure. In the context of the clause, apposition may interrupt the relation between major constituents, such as between subject and main verb. The latter is illustrated in examples (8) to (15). It may also interrupt the relation between clauses. Thus examples (16) to (19) illustrate how an adjunct clause is separated from the main clause which it modifies. Moreover, examples (20) to (22) show how co-ordinated (main) clauses may be separated by appositive structures. Perhaps the most important implication of all for syntactic complexity, however, arises from the potential of apposition to be used more than once in the sentence, either separately (as in example (14)) or consecutively (as in example (23)). As with the long subject construction, the excessive use of this interruptive function delays our comprehension of important items in the text. Finally, examples (24) to (27) have shown that apposition may occur across sentence boundaries. In such cases, there is a tendency towards marked structure in the second appositive. Also, the nature of the relation here is always made explicit by the use of adverbials which denote appositional meaning.⁵ We have noted, provisionally, that apposition across sentence boundaries can interrupt larger discourse structures. The latter will be discussed more fully in chapter four.

The examples of apposition that we have analysed enable us to appreciate the role of apposition within the clause and to develop some idea of its interruptive range within text. The types of interruption that we have considered in relation to apposition certainly do not exhaust the full range of possibilities; nevertheless, they provide a sufficiently comprehensive picture to show the potential of apposition for creating syntactic complexity in text. Moreover, they have provided us with enough categories for us to be able to analyse certain examples of 'difficult' text. Consider, for instance, example (28) below:

(28) ^S ^A ^V ^O ^M ^H ^Q
 [[Berger et al.] also consider more [direct conjunctions] between
^M ^H
 the two cognitive styles, for example, between the [componentiality
^Q ^C ^O ^M ^H ^Q
 of technological consciousness] (reality can be divided into
^C ^O ^M ^H ^Q
 self-contained units) and [the 'taxonomic propensity' of
^M ^H ^Q
 bureaucratic consciousness] (the urge to categorically classify
^O ^M ^H
 phenomena; and also [more basic disjunctions] such as between
 the technological emphasis upon the separability of means and ends]
 (the cogs the workers produce can go into a motor car as a
^C
 nuclear weapon) and [the bureaucratic 'non' separability of means
 and ends.]] (It's rarely a question of an individual just getting
 a form, it's essential that the correct procedure is gone through
 to receive it, for example, by application to the right place - the
 link between means and ends has a 'positive moral value.')

(D 305, Block 9, 16-17, p.26).

After a first reading of example (28) one's initial reaction is simply to consider the sentence as long and unwieldy. However, on closer examination the structure of the sentence is seen to consist, basically, of a S V O structure, where the transitive verb consider takes three nominal groups as objects. The overall grammatical structure of the sentence is therefore quite simple. It is when we consider the structure of its constituent

nominal groups that the reason for the initial difficulty caused by this sentence can be seen. The complexity of this sentence arises from an over-frequent use of apposition within the structure of the nominal groups. In the first nominal group, for example, there is an instance of double apposition (shown by the double lines). Here, the head (conjunctions) of the first member, together with its modification, is replaced in the second appositive by the head componentiality. The appositive relation between the two heads is made explicit by the adverbial for example. The head componentiality (and its modification) is itself replaced in the third appositive by the head reality. The third appositive is a passive clause, which consists of a subject, verb and object complement. In the second object nominal group, the head propensity is replaced in the second member of the relation by the head urge. There is no explicit item of appositional meaning for this relation. In the third, and final, object nominal group, the head disjunctions is replaced by the head emphasis and its modification in the following appositive. The syntactic complexity of this sentence therefore arises from the appositional complexity of its individual nominal groups. We might also note that the first apposition relation (i.e. double apposition relation) interrupts the relation between two co-ordinated nominal groups (signalled by and). Finally, there is also a long sentence-final 'comment' (contained within round brackets) which, while not making the parent sentence anymore complex, interrupts its relation with its adjoining sentence(s). The latter contributes to syntactic complexity at a more macro-level of structure (see chapter four).

A similar type of complexity is illustrated by example (29):

(29) ^A [Thus ^S [the journalistic attitude towards problem-solving,] ^A [apart from displaying the extrinsic occupational features of polarisation, dramatisation, etc.,] ^A also shows [the presence of exported modes of consciousness ^{O_M} [in its tendency to break down ^H [complex ^M processes] ^M (e.g. racial ^M integration, economic ^M recession) into simple quantifiable items ^H (the number of blacks in the U.K., the size of the gold reserves);]] [in its emphasis upon straightforward

casual relations (cutback ^M ^H immigration - hold wages at a certain

level);] ^A [in its concentration upon the allocation of the problem
to a specific agency which will then solve it through an appropriate
programme;] ^A [in its insistence upon avoiding 'extremes' of emotional
commitment.]]

(D 305, Block 9, 16-17).

Although this sentence is abnormally long, its complexity cannot be explained simply in terms of this. In one important respect, for example, its length is a necessary (and quite valid) consequence of the author's attempt at a rhetorical marking of the subject-matter. The latter is indicated in the strong repetition semantics of the four adjunct clauses, which each begin with in its. In fact, this structural and semantic repetition in some ways mitigates the sentence's abnormal length, creating in the reader a strong expectation of what is to follow. However, within the first adjunct clause there occur two instances of apposition. Both relations are of the instantiation-type; more specifically, their form is of the general-particular relation, whose form is also made explicit by such adverbials as for example, etc. For instance, the lexical head of the first appositive in each case is an abstract noun. The nominals which follow in the second appositives provide concrete particulars for the preceding generalization. Similarly, in the following adjunct clause there is another apposition relation, also of instantiation; this, too, could be made explicit by for example. In sum, there are thus three separate sets of apposition relations within a relatively short distance of each other in the same sentence. The first instance of apposition interrupts the relation between the direct object (complex processes) and the object complement (simple quantifiable items). Once again it is suggested that the syntactic complexity of this sentence is a function of its appositional structures.

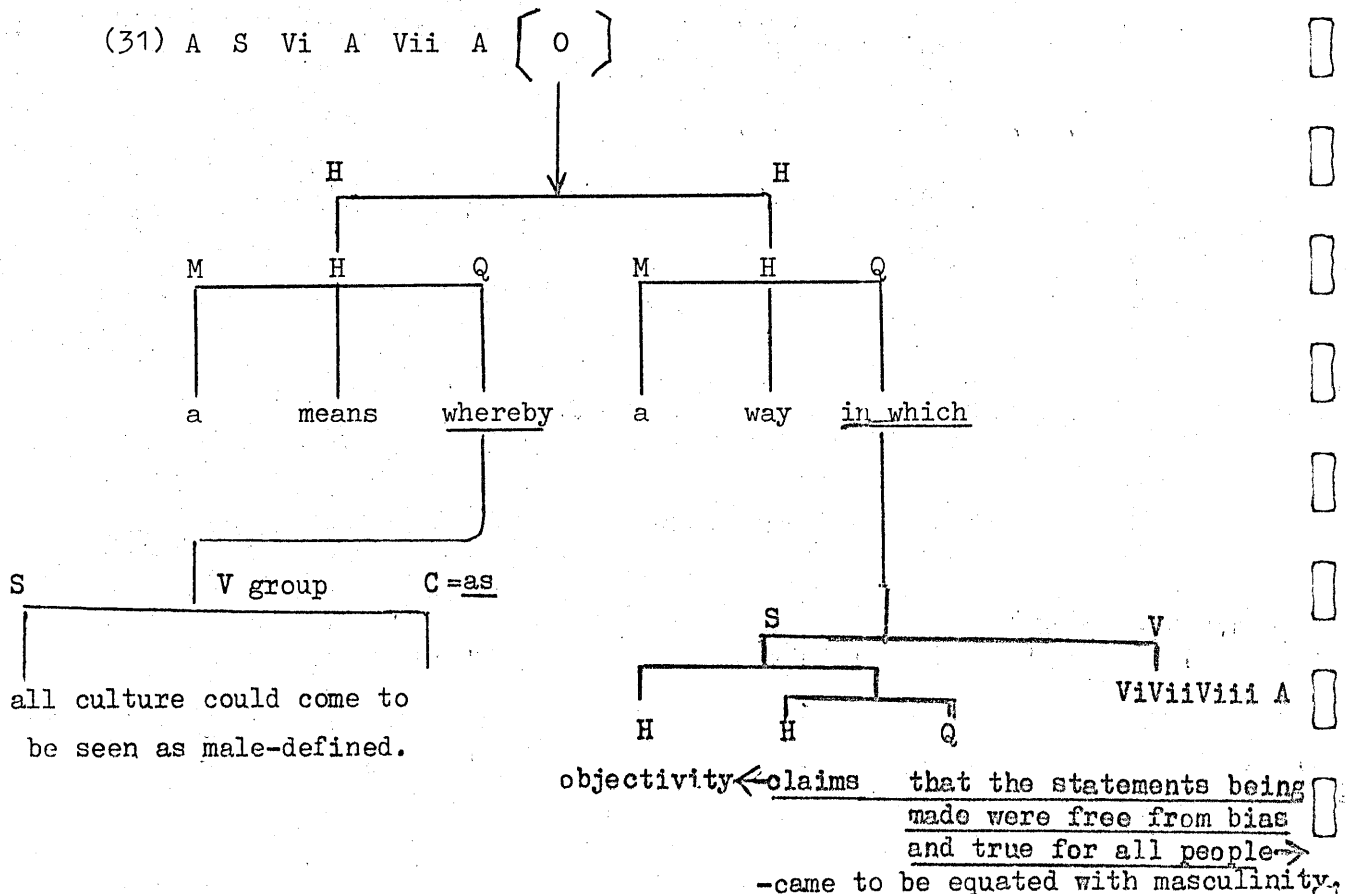
A comparatively less complex sentence is provided by example (30):

(30) ^A [In the nineteenth century,] ^S the sociologist George Simmel ^{Vi} was
^A already detecting ^{Vii} ^A [within the identification of 'man' with 'human
^H ^Q being',] ^{OM} [a means whereby all culture could come to be seen as

male-defined] [a way in which objectivity - claims that
the statements being made were free from bias and true for all
people - came to be equated with masculinity.]

(D 305, Block 9, 16-17, p.9).

The initial difficulty caused by this sentence arises from the interruption to the structure of the clause by the apposition relation. Basically, the sentence consists of an adjunct, subject, verb phrase and two objects; it is therefore quite uncomplicated in its overall grammatical structure. However, the second object nominal group of the transitive verb detecting is interrupted by the apposition relation. More specifically, the subject noun head of the postmodifying clause of the head objectivity is separated from its verb phrase by the second appositive. The precise nature of this interruption within the structure of the clause is made more clear by the following tree:



Example (31) shows the depth of the interruption by the apposition. In the apposed clause, the first head (objectivity) is replaced by the head claims and its postmodifying clause. Thus, although there is only one instance of apposition within this sentence, the sentence is quite complex because of the delaying effect of its appositional structure.

In conclusion, the reasons for analysing apposition in such detail in this section are two-fold. Firstly, we have attempted to show the important interruptive (and hence, 'delaying') effects of apposition within the structure of the clause. Considered from this point of view, apposition has much in common with the long subject construction (see chapter one). Secondly, we have seen that what might be referred to as the 'mechanics' of apposition consists, essentially, of replacing constituents from the first appositive structure. Such replacement focuses on 'heads' within the clause, which it may delete or lexically repeat. Like the modifiers which accompany the head in the nominal group, the second appositive is subordinate to the head which it postmodifies. The important implication here is that the use of apposition within the clause is a function of clarity of syntax. The latter proscribes the use of excessive apposition which either obscures the relation between lexical heads or interrupts crucial grammatical relations. It also proscribes the use of multiple apposition within the same sentence which delays the sentence boundary too long. In addition, apposition which occurs in many sentences in close succession can place a great load on the reader's tolerance; this, too, is to be avoided for a clear syntax.

I hope to have shown that the severity of the interruption by apposition depends on its position within the sentence. Thus, apposition need not be abnormally long to confuse the reader or delay his comprehension of the whole sentence. For instance, consider the following extract from a recent newspaper article:

(32) ^A ^S ^H
 [[As China warns the world that [the treaty following Salt II] -
^M ^H ^{Vi} ^A
 the second round of Strategic Arms Limitation Talks - will merely
^{Vii} ^O ^S
 free [Russia's hand for more mischief in the Far East,]] [Moscow
^{Vi} ^{Vii}
 and Hanoi] are admitting to [questionable military moves in Indo-
 China that they previously indignantly denied.]]
 (Observer, 3.6.79).

This sentence begins with an adjunct clause, followed by a main clause (beginning with the subject 'Moscow and Hanoi'). As can be seen, however, the adjunct itself is interrupted by an appositive nominal group. Here, the head Salt II is replaced in a rephrasing appositional relation. The interruption occurs between the subject of the embedded sentence of the matrix-clause (i.e. 'the treaty following Salt II') and its verb phrase. This sentence does not merely illustrate interruption between subject and verb, however. The difficulty of this sentence is a function of the interruption between subject and main verb in the environment of an adjunct clause at the beginning of the sentence. This lends weight to the point made earlier about the importance of the part of clause structure where apposition is allowed to occur. The advisability of using apposition thus depends not only on what constituent(s) it interrupts but on the place in the sentence where this interruption occurs.

SUMMARY

In this chapter we have explored the structure and function of the nominal group within the context of the clause (see especially section 1.1). Next to the main verb of the sentence, the lexical heads of the sentence are the most important items for comprehension. The various modifying structures which accompany the head of the nominal group therefore occupy a relatively less important role within the basic semantics of the sentence. If there is any question of syntactic complexity resulting from complexity within the structure of nominal groups, then the modifying structures which accompany the heads must be in some way responsible for this. Thus, in any attempt to rewrite such a complex sentence, the modifiers themselves should be the first items affected. This analysis of the structure of nominal groups has also explained in part the structure and function of apposition within the sentence. Apposition, too, relies on lexical heads within the clause and frequently consists of nominal groups. Like Q of the nominal group, the function of the second appositive is to postmodify the head of the first appositive. More specifically, the scope of backward reference of the second appositive extends to the (nearest) preceding 'similar' grammatical or lexical structure.

Chapter Three: Interpolation

Section 3.1: General Criteria

In the second section of the last chapter we introduced the notion of interruption to the structure of the clause by the postmodification (apposition) of the heads within the nominal group. Such appositive structures were seen to be extraneous items as far as the essential construction of the sentence is concerned. Like the normal modifiers M and Q of the nominal group, appositive structures can often be deleted from the sentence without making the sentence ungrammatical. In this chapter we will consider another structure which, like apposition, modifies different parts of the clause or even the whole clause itself. We will refer to this structure as interpolation.¹ As for apposition, interpolation is not an obligatory part of minimal clause structure but occupies a subordinate role within the sentence. Unlike apposition, however, interpolation can have either anaphoric or cataphoric reference to the item it modifies. Moreover, it is not constrained to modify only similar grammatical items; it need not, in fact, bear any grammatical resemblance to the item it modifies. This is not to suggest that interpolation does not replace some part of the preceding sentence, but this is not its distinguishing feature. The distinguishing feature of interpolation is rather to be sought in its semantic function within the sentence. This semantic function is to comment on or evaluate some part of the adjoining clause structure. Also unlike apposition, it does not further specify or narrow-down the meaning of some other item within the clause. As an essentially modifying structure interpolation can delay or interrupt major constituents of clause structure. When used excessively, it can delay our comprehension of the whole sentence or obscure the relation between important parts of grammatical structure; it may, also, delay the sentence boundary.

In the same way as apposition, interpolation is typographically signalled by a variety of punctuation marks. These punctuation marks comprise the following: the comma, the dash and (round) parentheses. In the case of the dash, there are two possible forms it may take to signal an interpolation. Firstly, if the interpolation occurs in the middle of clause structure, then two dashes are used (one preceding and one following) to demark the extent of the interpolation. Secondly, if the interpolation occurs in sentence-final position, then only one (preceding) dash is used and the interpolation is followed by the period. This situation is the same for

the comma. Sentence-final interpolation is interesting in that it contains two different forms of punctuation, the dash (or comma) and the full-stop. In other cases of interpolation (i.e. when not in sentence-final position) the interpolation combines the same forms of punctuation to indicate its boundaries. On this type of use of the dash Quirk (et al., 1972) remarks that it tends to "give a somewhat... dramatic and informal impression, suggesting an impromptu aside rather than a planned inclusion." (p. 1071.) Quirk maintains that such punctuation marks as the ones we have mentioned (i.e. the bracket, dash and comma) are not interchangeable. At the bottom end of the scale, the comma is the least obtrusive. However, when there is a possibility for confusion with neighbouring commas, brackets and dashes are preferred.

The type of interruption that Quirk seems to have in mind should not be identified with the grammatical type of interruption with which we are concerned. Quirk's notion of interruption has to do with levels of formality of the intrusion of additional information into sentence structure. Moreover, it does not seem correct to assert, as Quirk does, that interpolations denoted by the dash are not as planned as interpolations indicated by other means (i.e. commas). As we hope to show in what follows, if we recognise that interpolation performs a well-defined and specific semantic function within the clause, then its use in text is both purposeful and deliberate (notwithstanding the punctuation used). Further, it is argued here that the criteria for recognising interpolation are semantic rather than typographical. In respect of the purely grammatical type of interpolation we are discussing, any consistent use of the types of punctuation mentioned is acceptable for interpolation. The word 'consistent' is crucial here. From the point of view of consistency, the punctuation in example (1) below is surely wrong:

S V C

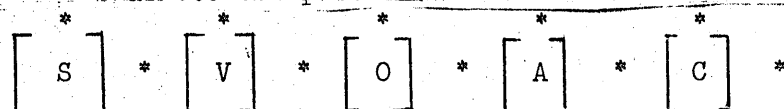
- (1) He was [not only a great reporting cameraman] - which his awards recognised (they included Cameraman of the Year for his portfolio of Vietnam films with Julian Pettifer), he was also a remarkable loner, a man who believed in flying his own plane anywhere in search of a visual story.

(The Listener, 15.3.79).

The total extent of the interpolation in this example is indicated by the single line and is used to comment on the complement of the preceding clause 'not only a great reporting cameraman'. The first (superordinate) interpolation

embeds another interpolation, which is denoted by brackets. Because the inclusive interpolation begins with a dash, it should also be followed by a dash after the second bracket. To have a comma after the second bracket is therefore incorrect here. Apposition can make use of any of the punctuation used by interpolation, especially the dash. Thus, the semantic criteria for interpolation are important because they serve to differentiate it from apposition. The main point about the punctuation used by both apposition and interpolation, however, is that it signals the interruptive nature of these two different forms of modification.

The distribution range of interpolation is equal to that of apposition. For example, it can interrupt the boundaries between major constituents or it can interrupt the structure of the constituents (see P.130⁹) themselves. The following diagram makes this point more clearly, where the asterisks denote the possible areas of interruption:



In the diagram V denotes verb phrase. Obviously, this level of generality does not capture the full range of interruption by interpolation. For example, it does not show what part of the constituent may be interrupted. In the second section we will present a more comprehensive picture of this range through examples. An important point illustrated by the diagram is that interpolation never occurs in sentence-initial position. Thus, if we were to select an interpolation which occurred in any of the asterisked positions in the diagram and place it at the beginning of the sentence, this same structure would be no longer qualify as an interpolation. The latter is the only distributional constraint on the use of interpolation within the sentence. Finally, the interpolated item need in no way resemble the part of structure which it interrupts. There is, also, no typical grammatical forms of interpolation; it can vary from a nominal group, adjectival or verbal group to a whole clause or sequence of clauses. An interpolated clause need not, however, contain any verb; it may be a verbless clause.

So far, we have said that the characteristic semantic function of interpolation is to comment on or evaluate some part of the clause. In other respects, however, the semantic character of different grammatical forms of interpolation is not as uniform as this might suggest. This difference is realised once we consider interpolation in the context of the clause; more precisely, when we consider the information-value of the clause (for subordinate ~~and~~

and main clauses). The two possible types of interpolated clause, subordinate and main, perform quite different contextual roles within text. The contextual function of the subordinate clause, for example, is to present 'given' information (theme), while that of the independent clause is to present 'new' information (rheme). In the case of the former, the information presented is contextually-dependent; in the latter, the information is relatively independent of context.

As a concrete illustration of the contextual difference between given and new information in the clause, we may compare the non-restrictive relative clause with its corresponding independent form. Both types of clause are possible forms of interpolation. Consider, for instance, examples (2) and (3) below:

- (2) ^S Mrs Smith - ^{Vi} and she should know better - ^{Vii} has approached ^O [that awful man in the ministry.]
- (3) ^S Mrs Smith, ^{Vi} who should know better, ^{Vii} has approached ^O [that awful man in the ministry.]

In each of these examples, the interpolated clause is underlined. In example (2), the independent clause is demarked by correlative dashes. The interpolation adds new information to the second independent clause, which has the structure S Vi Vii O (shown by the superscripts). The grammatical indicator of the independent status of the interpolated clause (apart from its internal grammatical structure) is to be found in the interpolating use of and. It is a rule of the syntax of English, for example, that co-ordination combines similar grammatical structures. In short, example (2) consists of two independent clauses. In contrast, example (3) consists of an independent clause (also with a structure of S Vi Vii O) and an interpolated subordinate clause, which is denoted by correlative commas. While the independent clause of example (3) provides the 'new' information of the sentence, the non-restrictive relative clause contextualises this information in terms of what is already known about the subject, Mrs Smith. The basis for the latter assumption is that the information provided by the subordinate clause has either been mentioned in some part of the preceding discourse or simply because the speaker or writer assumes this knowledge on the part of the hearer or reader. Both the non-restrictive relative clause and the independent (declarative) clause receive separate tone unit boundaries from the other clauses in the sentence when uttered in actual

speech. The independent (declarative) clause would receive a falling intonation nucleus, while the non-restrictive relative clause would receive a fall-rise. This difference is illustrated by examples (4) and (5) below, where the individual tone groups are denoted by vertical lines:

- (4) | ^VThe teachers - | they are communists - | are ready to strike. |
- (5) | ^VThe teachers, | ^Vwho are communists, | are ready to strike. |

The different contextual roles performed by the subordinate and main clause lies at the basis of cohesion in text (Winter 74:501 and Winter 77: 45/52). The information structure in text, for instance, depends for its coherence on the alternation between given and new information. After it is first introduced, the new information provided by the interpolated independent clause becomes part of the given information of the text. This chaining effect may in fact be taken as part of the explanation for the ability we have to differentiate between 'connected' and 'unconnected' discourse (Keen, 1978). In sum, both the subordinate clause and the main clause are valid vehicles for interpolation meaning, but they play quite different contextual roles in text.

It is convenient at this point in our discussion to consider the analysis of interpolation within current transformational-generative grammar (henceforth T.G.). In what follows, it will be argued that T.G. does not distinguish between relativisation and interpolation in terms of their different contextual roles within text.² One possible reason for this is that T.G. is almost exclusively concerned with specifying grammatical relationships within the same sentence. But we have just seen in relation to the non-restrictive relative clause that grammatical relationships also obtain between sentences. It follows that such inter-sentential relations cannot be satisfactorily explained by confining attention to the sentence alone. It is a serious weakness of T.G. that it does not provide for an adequate account of linguistic context in the sense that we have been discussing (van Dijk, 1973, p.37). We may begin by examining the analysis given in B. Jacobson's book 'Transformational Generative Grammar' (North-Holland, 1977), p.341, of non-restrictive relative clauses.

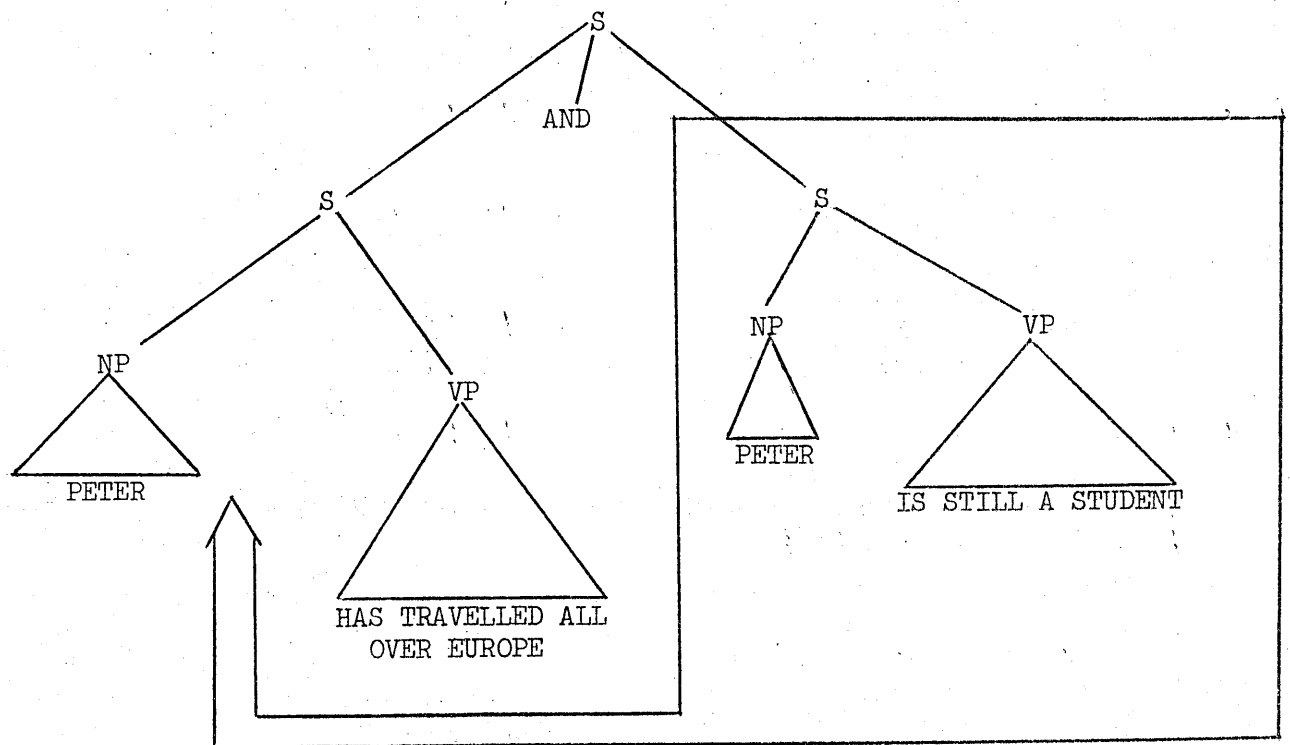
According to Jacobson's account, for example, non-restrictive relative clauses are derived from two conjoined sentences. The non-restrictive relative clause cited by Jacobson is given below in example (6):

- (6) Peter, who is still a student, mind you, has travelled all over Europe.

This sentence is assumed to be derived from two conjoined sentences, given in examples (7) and (8):

- (7) Peter has travelled all over Europe, and Peter is still a student, mind you.
 (8) Peter - and Peter is still a student, mind you - has travelled all over the Europe.

The derivation of this sentence is shown by the following phrase-marker (omitting 'mind you'):



At this stage, two transformations are possible. In the first one, 'and' is dropped and the second conjunct is adjoined to the right of the subject NP of the first conjunct. After this, a relative rule is applied which converts the second NP to the corresponding pronoun, which will derive

sentence (6) above. The alternative transformation retains 'and' and does not adjoin the second conjunct to the right of the subject NP of the first conjunct. This latter transformation would yield sentence (8) above. The operation of both T. rules is clear and straightforward. However, the operation assumes that both derived structures are equivalent in meaning because they have the same common base. That is, both the non-restrictive relative clause in sentence (6) and the interpolated clause in sentence (8) are assumed to be grammatically equivalent transforms of sentence (7). But we have seen that both types of clause (i.e. subordinate and main) are quite distinct in terms of their contextual meaning: while the one presents given information, the other presents new information. The two types of clause cannot therefore be equivalent in meaning and hence, the transformations posited here cannot be meaning-preserving (Partee, 1971). T.G. suffers from an excessive concern with intra-sentential relations. As a consequence of this, it cannot provide an adequate account of sentence meaning which, as we have seen, is related to a wider linguistic context.

This discussion of the role of subordinate and independent clauses as interpolative items within clause structure has important implications for syntactic complexity within the clause. For instance, if we accept that the function of the independent (declarative) clause is to present 'new' information, then there is a possibility of over-loading the information-structure of a sentence if it is interrupted by many such independent clauses as interpolations. Syntactic complexity would also result if the sentence were interrupted by an excessively long interpolated independent clause. Winter (forthcoming) calls the clause containing the interpolated clause a host-clause. Because of the adjunct-like mobility of interpolation, a great variety of interruptions of the same clause is entirely possible. However, clarity of syntax places severe constraints on the number of such independent clauses within the same sentence. For example, we would not seriously expect to encounter more than one independent clause interpolated within a host (declarative) clause. This notion of interruption illustrates a confluence of syntax and semantics. Thus, while syntactic complexity could be caused by the nature and extent of the interruption, the interruption itself would have implications for the semantic complexity of the sentence as well.

At the other extreme, the subordinate clause, when it occurs as an interpolation, may also contribute to syntactic complexity for semantic reasons. Subordinate clauses are notoriously more difficult to process

than their corresponding independent forms. There are two possible explanations for this, both of which are semantic in character. According to the first explanation, the contextually-dependent nature of subordination might well entail more sophisticated semantic processing than corresponding independent clauses. An alternative syntactic explanation is that the syntactic relationships introduced by subordination are more complex than the relationships entailed by the independent clause (Morgan, 1978, p.4). Research into subordination has pointed to the complex dependencies created by this type of clause. Goodman and Greene (1977) claim that the relative difficulty of subordination is especially marked for young children. In relation to such subordinating conjunctions as although, even though, unless or when, they state that "even when children can read these structures smoothly, they provide evidence when they discuss or retell what they have read that they do not understand the semantic relationships." (P.24). The authors claim that this is due to the fact that co-ordinating conjunctions are structurally more flexible and semantically less constrained than subordinating conjunctions. According to their argument, co-ordinated structures are principally concerned with consecutive or concurrent temporal relationships. With subordinate structures, in contrast, we introduce the more subtle semantic factors of 'cause' and 'effect'. However, one problem with this argument is that, in addition to expressing temporal relations and evaluations, co-ordinating structures can also express deductive inference. This is illustrated by example (9) below:

(9) It was raining so we sheltered in the house.

The distinction made by the authors between the semantic relations entailed by subordination and co-ordination therefore seems rather dubious. At least, their claim remains to be proved. In sum, the former semantic explanation for the relative difficulty of subordination seems preferable although this, too, remains to be demonstrated.

In conclusion, we may make the following generalisations about the use of interpolation to interrupt the structure of the clause. We have already referred to the rich syntactic mobility of interpolation within the clause. Because of this mobility, interpolation should be seen as a special form of adjunct function within the clause. In broad terms, adjuncts are those clause elements other than pre- or postmodifiers which can be removed from the clause, leaving the basic structure S V (O)(C). In a sentence of S V structure

we would expect to find an adjunct in any of the following positions (denoted by asterisks):

* S * V *

In a sentence of S V C O structure, adjuncts could occur in any of the following positions:

* S * V * C * O *

The important restriction on the mobility of interpolation, however, is that it cannot (unlike the ordinary adjunct) occur before the subject itself. This restriction has to do with the semantic function of interpolation within the sentence. Its characteristic feature is to comment on or evaluate in some way some part of the preceding or following clause structure. Clearly, the great advantage of interpolation for the writer is that it provides a vehicle for his subjective judgement, which he can use to modify anything he says.³ Insofar as it is also an extraneous postmodifying component of the sentence, interpolation is very similar to apposition.

As we will see in section 3.2 of this chapter, interpolation may take a variety of grammatical forms. Of the various syntactic forms it may assume, it has been suggested that its status as an interruptive clause (independent or dependent) could create the most serious problems for sentence clarity if used excessively. On the other hand, there is scope for syntactic complexity when a host (independent) clause is interrupted by another independent clause or clauses. It is suggested that the latter could over-load the information-structure of the sentence. On the other hand, because the information presented by the subordinate clause is context-dependent, it introduces additional semantic relationships into the sentence. Thus, if it occurs more than once (as an interpolation) within the same clause, the reader may be confused by the additional semantic relations introduced. Finally, because of the quite different contextual roles performed by the subordinate clause and independent clause, it is a serious weakness of T.G. that it confuses this distinction.

Section 3.2: Examples of Interpolation

In this section an attempt will be made to exemplify the points made in section 3.1. Firstly, we will present a short collection of excerpts from texts to give some idea of the distributional range of interpolation. Secondly, we will then present a much more detailed and extensive analysis of interpolation. A great many of the examples here will be culled, once again, from Open University course units. A slightly larger proportion of the examples presented will be taken from journalist articles. The purpose of the analysis will be to illustrate the semantic function of interpolation within the sentence. That is, both its replacement meaning and evaluative function will be considered in relation to the items of the clause which it interrupts. Finally, in the light of this discussion, we will consider an example of interpolation which seems to create unnecessary syntactic complexity.

A brief indication of the mobility of interpolation throughout the sentence can be given by listing the most typical contexts for its occurrence. In the examples which follow, the dotted lines indicate the interrupted structures, while the single horizontal lines denote the interpolated structures. In this instance, consider the following:

- (1) Interpolation used to interrupt a noun head and its premodification:

The use of an all - in - step - or, as it is more usually described, 'coherent' - radiation to carry information can be effected in principle at any wavelength.

(Times, 28.1.66).

But while we ought to repudiate the illusory - and, if translated into nuclear terms, dangerous - independence from America, we should also recognise that the criticism that Britain is trying to live politically above economic means carries the sting of truth.

(Observer, 29.1.67).

- (2) to interrupt the verb and its object (in this case a noun clause):

They suggest in all seriousness - and, what is even more remarkable, the suggestion has been taken seriously - that in the 1914-18 war 10

Bavarian officers, all Protestants, served with Roessler in the same company

(Observer, 8.1.67).

- (3) to interrupt the subject and main verb:

The theory - and it is only a theory - is that those early frogs fed on smaller animals along the shores and banks of ponds and streams, and the shortest cut to safety when attacked was to leap into the water.

(New scientist, 23.6.66).

This conclusion - and it is not just Mr. Powell's conclusion - is that two of their assumptions are no longer valid.

(Guardian, 26.10.65).

The only real alternative - and the survey's forecasts make this point clear - is more road-building in London.

(Observer, 16.1.66).

- (4) to interrupt the main verb and adverbial adjunct:

They operate, of course - and so they should - under stringent checks...

(Observer, 7.10.66).

- (5) to interrupt a subordinate clause and a main clause:

Now with the new law rigorously enforced - this year there are spot checks all along the route - the character of the Tow is likely to change.

If their prestige is committed to sustain communism in Cuba - especially now that China can now claim no part of it - CASTRO shows that he believes his own revolutionary status in Latin America depends on keeping up the "implacable fight" against the United States.

(Daily Telegraph, 31.8.66).

- (6) to interrupt an adverbial adjunct from a main clause:

With difficulty - it is an open secret - Mr. Robinson has persuaded his cabinet colleagues to continue talks...

(Guardian, 10.1.66).

The examples given above give some idea of the interruptive range of interpolation. We can now begin to analyse these principal types of interruption in more detail. Consider, for instance, example (7) below, where the interpolation interrupts the relation between subject and main verb:

- (7) [[^S ^M A new English Development ^H Agency] - the ^H proposal which the Government seems most likely to adopt- ^A ^{Vi} ^{Vii} simply means setting up ^O another public body. funded with public money, to make investments, some of which are too risky for the banks to back.]]

(Guardian, 16.3.79).

In this example, as in all those which follow in this section, the various superscripts and brackets denote the sentences constituency-structure. As with examples (1) to (6), the broken lines indicate the part of structure interrupted, while the single horizontal line denotes the interpolation itself. In example (7), there is a clear case of replacement between two nominal groups. The subject nominal group 'A new English Development Agency' is replaced by an interpolated nominal group, which consists of a noun head (proposal) and a restrictive relative clause ('which the government seems most likely to adopt'). This restrictive relative clause supplies the actual comment of the interpolation.

In the second sentence of example (8) the relation between subject and main verb is also interrupted:

- (8) (1) Why have tenants not bought before?
(2) [^A ^{A'} Probably [because [a modern council flat built to Parker Morris standards] (often higher than those of private estate housing) is [a snip at a rent around one tenth of national average earnings.]]]

^A → (The Daily Telegraph, 27.3.79).

In sentence (2), the subject is the subject of an adjunct clause. It is

interrupted from the main verb (is) by a comparative clause ('often higher than those of private estate housing'), typographically signalled by round brackets. The interpolated clause evaluates the prepositional complement 'Parker Morris standards' included within the subject.

In the interruption of subject and verb in example (9), the subject is the subject of the embedded sentence of a matrix-clause:

- (9) $\begin{matrix} A \\ \left[\left[\text{Whatever the truth of specific allegations,} \right] \left[\begin{matrix} S_1 & V & C \\ \text{it appears most} \\ A \end{matrix} \right. \right. \end{matrix}$
 $\begin{matrix} S_2 \\ \text{likely that} \left[\begin{matrix} C & A \\ \text{the Ministry of Information} \end{matrix} \right] - \text{under the superintendence} \\ C & A \end{matrix}$
of Dr. Mulder, and probably with the connivance of some in the
 $\begin{matrix} V & O \\ \text{Government} - \text{waged} \left[\text{a secret propaganda war with public funds.} \right] \end{matrix}$

(The Daily Telegraph, 27.3.79).

In broad terms, the total sentence is of the structure A S V O. The sentence also illustrates extraposition of the subject (see section 1.3 of chapter one). The dummy-it is denoted by S_1 , while the extraposed subject is shown as S_2 . The interpolation itself consists of two co-ordinated (commenting) adjuncts which postmodify the extraposed subject.

In example (10), the subject, which is interrupted from the main verb, is preceded by a long composite adjunct:

- (10) $\begin{matrix} A \\ \left[\left[\text{In the long history of cynical and manipulative decisions taken} \right. \right. \end{matrix}$
 $\begin{matrix} \text{by the United States in the course of its frequently conflictive} \\ \text{relationship with the countries of Latin America over a century} \\ \text{and a half,} \end{matrix}$
 $\begin{matrix} S \\ \left[\begin{matrix} \text{the speech on Thursday by Cyrus Vance} \end{matrix} \right] - \text{calling} \\ \text{for the downfall of President Somoza and his replacement by a} \\ \text{government of national reconciliation - will surely deserve} \left[\begin{matrix} Vi & A & Vii \\ \text{a} & & \text{prominent place.} \end{matrix} \right] \end{matrix}$

(Guardian, 23.6.79).

The interruption in this example is between the subject 'the speech on Thursday by Cyrus Vance' and its verb phrase 'will surely deserve'. The interpolation consists of a present participle clause, which is used to comment on and further specify the meaning of the noun head speech. We may also note, in passing, the frequent propositions of the adjunct clause which precedes the subject of the sentence. These prepositions are indicated

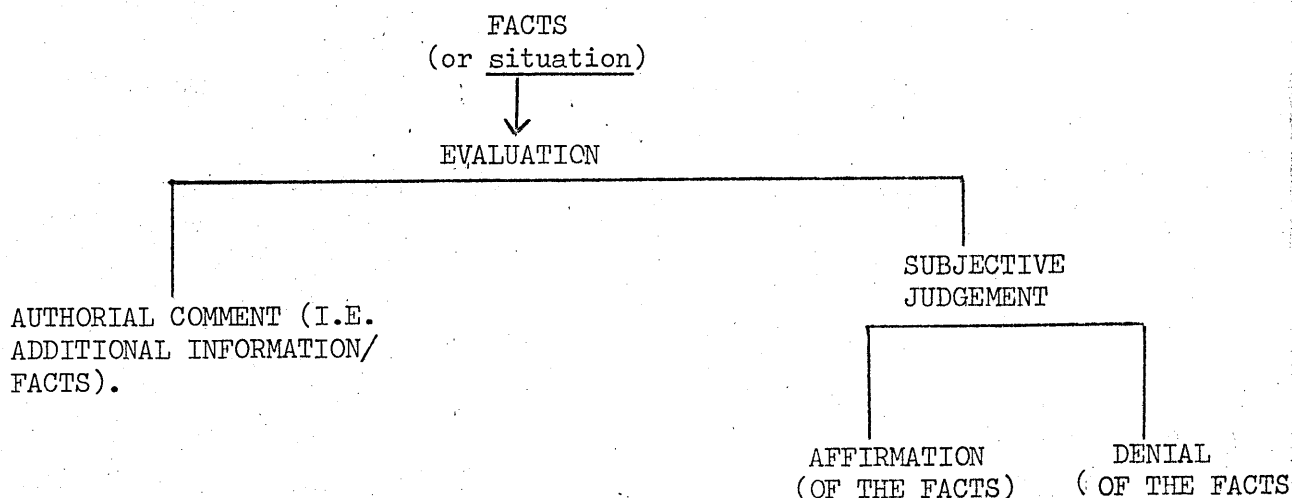
by the short horizontal lines. Such a frequency of prepositions within a single (sentence-initial) adjunct clause is best avoided in the interests of clarity of syntax.

As an instance of the interruption between the verb phrase and its adjunct, consider example (11) below:

(11) S Vi Vii Vii A Vi Vii Viii A
 [[Discrepancies] could be found (indeed have been found) among
other groups of cultures.]]

(D 305, Block 8, 14-15, p.81).

The verb phrase in example (11) consists of a modal auxiliary (could), the auxiliary verb be and the past participle verb found. The aspect of the verb phrase is therefore possible event. This verb phrase is replaced by a present perfect verb phrase which affirms the truth of the possibility. This affirmation is made explicit by the adjunct indeed at the beginning of the interpolation. This example is also interesting in that it illustrates another aspect of the evaluative function of interpolation. Evaluation, for instance, is merely a generic term for the subjective judgements of the writer (or speaker). It subsumes authorial comment, affirmation of the truth of a preceding clause or structure, or a denial of the truth of something in the preceding text. Evaluation may therefore have either a positive or a negative orientation. If we assume the statement preceding the interpolation as representing the facts of a situation, then the interpolation will represent the subjective assessment of the writer in the light of these facts. That is, whether he considers the facts to be true of false, or whether he needs to add more specific information (commentary) to these facts. This idea may be made more clear if we represent it in the following way:



In example (11), the verb phrase which is interpolated is thus an aspect of subjective judgement on the part of the writer.

The verb phrase itself can be interrupted by interpolation. This is illustrated sentence (2) of example (12):

- (12) (1) This block will follow William James in stressing the nature of subjective experience...
- (2) [But it will go beyond William James [^{Vi}by attempting -
^Aalbeit [^Ain an introductory fashion] - ^{Vii}to link [^Osubjective
experience with objective reality,] ^Cand to ^Vshow [^Othe ways
in which being and acting within the world ^Vinfluence (or
even ^Vdetermine) [^Oour conception, not only of that world,
but of ourselves.]]]]

(D 305, Block 9, 16-17, p.8).

In this example, the present participle attempting is interrupted from the infinitive verb to link by an interpolation which consists of two adjuncts. The semantic function of the interpolation is that of the simple adverbial function of postmodifying the verb attempting. There is also a second instance of interpolation in this sentence. The latter interrupts the transitive verb influence from its object. This interpolation replaces the verb influence with the transitive verb determine. This second instance of interpolation is embedded within the object of the verb to show; that is, the second object in this sentence embeds another object.

The nominal group is open to interruption by a variety of interpolative forms. In this instance, consider example (13):

- (13) [[^HIf the occupants of the residences can be identified from ^Mother
^Msources] - the obvious ^H[one] ^Vbeing census returns - ^Cgiving their
^Moccupation, and perhaps their rateable ^H[value] ^Q(from ratebooks),
^Qsize of family, employment of servants, place of birth, etc.,] [then
it can be ascertained whether, and to what extent, social polarisation
on a spatial plane existed.]]

(D 301, 13-15, p.32).

In this example, the sentence begins with a conditional if-clause, which contains a prepositional complement in the form of the nominal group other sources. The head sources is replaced by the head one in the interpolated structure, which consists of the structure S V C. The effect of this interpolation is to interrupt the relation between the noun head sources and its postmodification (Q). The interpolation is clearly evaluative in its use of the attributive adjective obvious. This sentence also contains a second instance of interpolation (typographically signalled by round brackets) in the form of a postmodifier (prepositional phrase) to the noun head value. The semantic function of this latter interpolation is to comment by the addition of further information.

In example (14) below the noun head beliefs is also interrupted from its postmodification by the (evaluative) interpolation:

- (14) ^S Our inferences concerning the characteristics, qualities and inner states of others ^V are ^C based on prior beliefs (^A sometimes ^M quite erroneous) ^Q about how these features are expressed and associated with one another.

(D 305, Block 8, 14-15, p.54).

The evaluative nature of the interpolation in this example is indicated by the use of the postmodifier erroneous. The interpolation itself consists of an adjunct and two modifiers. Strictly speaking, the last two items of the interpolation consist of an intensifier and adjective. In contrast to examples (13) and (14), the interruption in the second sentence of example (15) occurs at the end of the premodification, thus separating the premodification from the noun head sense:

- (15) (1) Was KME doomed from the start because it was an uneconomic Government subsidised job-saving scheme. (2) [^H Was it ^S doomed ^M [because in a strict legal (and philosophical) ^C [sense] it was [not a co-operative at all?]] (KME never felt able, under current legislation, to register as a co-operative).]

(Guardian, 29.3.79).

The interpolation here consists of the conjunction and and the adjective philosophical. The total interruptive sequence occurs within the context of a subordinate clause, which is indicated by the subordinator because. We might recall in this connexion our suggestion earlier of the greater difficulty of subordination for the reader. If this suggestion is accepted, then excessive interpolation within the context of a subordinate clause should be avoided all the more. The sentence also contains an instance of sentence-final interpolation (see below for a fuller discussion). The

function of the interpolation here is to comment on the noun head co-operative in the preceding clause. While this type of interpolation does not interrupt the relation between the major constituents of the sentence in which it occurs, a consequence of its use is to delay the sentence-boundary. In this sense, it may be said to interrupt the relation between the sentence and its adjoining clause structure.

A frequent type of interruption by interpolation occurs between clauses. However, the clauses so interrupted are of various kinds. Consider, for instance, example (16) below:

- (16) ^S₁ ^V ^C ^S₂ [It is interesting to note that [in criminal law the crime is defined by the intention to commit it (hence the distinction between manslaughter and murder), whereas in civil law the offence is defined without reference to intention.]]]

(D 305, Block 8, 14-15, p.64).

In this sentence, the interpolation, which is signalled by round parentheses, comments on the embedded (independent) clause of the matrix-clause. The interpolation interrupts the relation between this embedded independent clause and its adjoining subordinate clause (which begins with the subordinator whereas). The semantic function of the interpolation is to introduce a deductive sequence into the sentence; this is made explicit by the adverbial item hence.

In example (17) below, the relation between a subordinate and main clause is also interrupted by the interpolation:

- (17) ^A ^{Vi} ^{Vii} ^H ^Q [Having lost [command of the commons] - ^M ^H the essential test of whether a Prime Minister goes or stays - ^S ^A he now faces ^O a spring election at a time he would never have chosen.]]

(Guardian, 29.3.79).

The subordinate clause in this example takes the form of a non-finite (present) participle clause. The meaning of the object of its verb lost ('command of the commons') is then evaluated by the interpolation. The interpolation replaces the head command by a new head (test) and its modification. The relation of evaluation is made explicit by the attributive adjective essential, which modifies the new head test in the interpolation. The interpolation itself consists of a nominal group which replaces the nominal group in the preceding participle clause, while the main clause following the interpolation is of the structure S A V O.

A main clause is also separated from a subordinate (adjunct) clause in example (18):

- (18) ^A[[And in the SALT dialogue]^S[[the Carter administration]^{Vi Vii} is taking
^Orisks with his own public opinion (though not with United States
Security) ^Aby offering the Russians a treaty in which they by no
means get the worst of the bargain.]

(Guardian, 29.3.79).

Basically, this sentence consists of the structure A S Vi Vii O A, and the interpolation occurs at the end of the object in this structure. More generally, the interpolation interrupts the relation between the main clause 'the Carter administration is taking risks with his own public opinion' and its adjoining adjunct clause 'by offering the Russians a treaty in which they by no means get the most of the bargain.' Once again, we have a subordinate clause ('though not with United States security') functioning as an interpolation. The semantic function of this interpolation is to postmodify the preceding main clause.

In a large number of cases, the clauses interrupted by interpolation are explicitly co-ordinated structures. Consider, for instance, example (19):

- (19) ^c[[And yet, somehow,^{Vii} [the liability for negligent statements]^c ^Ais
^Snot imposed [on barristers]^{Vi} ^Aand - in cases not cited in the
^{Vii}book (perhaps he would rather forget them) - [Lord Denning's^{Vi}
hospitality to the individual faced with group pressure]^{Viii} ^Ahas not
^{Oind.}been readily extended [to the student and the welfare recipient.]^c]

(Listener, 15.3.79).

So far, the examples of interpolation that we have considered have been anaphoric in reference. In contrast, example (19) contains an interpolation which is cataphoric in reference. The relation here must be cataphoric since the interpolation occurs after the conjunction and, which conjoins the preceding independent clause with the following independent clause. The interpolation here is especially striking in that it comprises a double-interpolation: one interpolation (signalled by round brackets) being embedded within the other. The second instance of interpolation refers back as a comment (of adding more information) to the first interpolation. However, taken together, the two interpolations clearly refer to the following main clause, whose structure is S Vi Vii O. More precisely, the

(D 301, 13-15, p.18).

Here, the interpolation evaluates the noun complement of the preceding independent clause as 'easier to undertake than dynamic analysis'. In the second co-ordinated structure of the sentence, the subject status is deleted. The sentence also contains two instances of apposition (signalled by round brackets). The second instance of apposition is a rephrasing relation, which is made explicit by the adverbial that is. This same relation also repeats the attributive adjective elusive in the second appositive. The two co-ordinated structures which are interrupted in this sentence are not identical in structure. The first, for example, is of the structure S V C, while the second is S V O. Examples (19) to (23), in sum, illustrate the interruption of conjoined independent structures by either anaphoric or cataphoric interpolations. They have also illustrated the deletion that may occur in the coordinated structures and that the coordinated structures need not be identical in grammatical structure. Perhaps a less frequent type of interpolation is that which separates the object complement from the direct object. This is illustrated in example (24) below:

- (24) $\left[\begin{array}{c} \text{Od.} \\ \text{What most people describe as the British Disease} \end{array} \right] \begin{array}{c} \text{S} \\ \text{I} \end{array} \begin{array}{c} \text{Vi} \\ \text{consider} \end{array} \begin{array}{c} \text{Vii} \\ \text{to be} \end{array} - \begin{array}{c} \text{A} \\ \text{in the long run} \end{array} - \begin{array}{c} \text{CO} \\ \text{one of Britain's strongest and} \\ \text{most attractive virtues.} \end{array} \right]$
- The Listener, 16.3.79).

The first thing we notice about this sentence is its marked grammar; that is, through the fronting of the direct object (WH-clause), the sentence highlights the semantic importance of this item. The interruption in this sentence is by an interpolated adjunct which separates the direct object of the verb consider and the subject and verb phrase from the object complement. The reference of the interpolated adjunct is thus cataphoric to the following object complement. The corresponding unmarked form of example (24) would be:

- (25) $\begin{array}{c} \text{S} \\ \text{I} \end{array} \begin{array}{c} \text{Vi} \\ \text{consider} \end{array} \begin{array}{c} \text{Od.} \\ \text{what most people describe as the British Disease} \end{array} \begin{array}{c} \text{Vii} \\ \text{to} \end{array} \begin{array}{c} \text{A} \\ \text{be} \end{array} - \begin{array}{c} \text{A} \\ \text{in the long run} \end{array} - \begin{array}{c} \text{Co.} \\ \text{one of Britain's strongest and most} \\ \text{attractive virtues.} \end{array} \right]$

We might note here that even in the unmarked sentence of example (25), the reference of the interpolation is still cataphoric.

In the first section of this chapter, reference was made to the non-restrictive relative clause as a possible form of interpolation. An illustration of this provided by example (26) below:

- (26) ^A
^S [Whenever the election] ^{Vi} (which is a large question in itself) ^{Vii C}
^S [the future of the EEC] will be [part of the campaign.]]

(Guardian, 21.3.79).

In this example, the interpolation (typographically signalled by round brackets) separates the adverbial adjunct from the main clause (of structure S Vi Vii C). The interpolated non-restrictive relative clause evaluates the preceding adverbial adjunct. The typographical signalling of the non-restrictive relative clause here is to be contrasted with that of example (27) below:

- (27) ^{Vi} [The Conservative Party] ^{Vii} is pledged to repeal ^{Viii} [the Education Act ^O
of 1976] which required local education authorities to submit
plans for going comprehensive - ^c and ^S [the message from Conservative
Central Office to the local parties] ^{V C} is clear : 'Hang on, help
is coming.']

(The Weekly Educational Review, 29.2.79).

In this example, correlative dashes are used to demark the extent of the interpolation. The interpolation here interrupts the relation between two independent clauses, the first of which has the structure S Vi Vii Viii, while the second has SVC. The non-restrictive relative clause interpolation postmodifies the direct object of the verb repeal in the preceding independent clause. As we saw in section 3.1 of this chapter, Quirk (et al., 1972) would argue that the typographical signalling of examples (26) and (27) is more informal and intrusive than the corresponding form where correlative commas are used. As far as the essential modifying function of interpolation is concerned, however, these three types of punctuation are interchangeable and do not change the meaning of the interpolation.

Thus far, the examples that we have considered have been of interpolations which occur in the middle of clause structure. We may now consider some examples of sentence - final interpolations. Such interpolations do not interrupt the structure of the clauses but instead delay the sentence-boundary of the sentences in which they occur. Consider, for instance, example (28):

(28) $\left[\begin{matrix} A & S \\ \text{So} & \left[\text{the industry's case} \right] \end{matrix} \right] \begin{matrix} V & C \\ \text{is} & \text{fine} \end{matrix} - \left[\begin{matrix} A \\ \emptyset \text{ as far as it goes.} \end{matrix} \right]$

(Guardian, 21.3.79).

In this example, the interpolation consists of a sentence-final adverbial adjunct which evaluates the scope of the preceding complement (predicative adjective) in the preceding independent clause, of structure A S V C. Once again, if we were to prepose this adjunct, it would be no more than an adverbial adjunct; it would no longer qualify as an interpolation. The replacement that occurs in the interpolation takes the form of a deletion of the preceding independent clause.

In the second sentence of example (29) the replacement is between nominal groups:

(29) (1) $\left[\begin{matrix} A \\ \left[\text{Over the years} \right] \end{matrix} \right] \left[\begin{matrix} S \\ \text{innumerable crusades,} \end{matrix} \right] \left[\begin{matrix} A \\ \text{running almost into} \end{matrix} \right]$
 $\left[\begin{matrix} \text{double figures,} \end{matrix} \right] \begin{matrix} \text{have taken} \end{matrix} \left[\begin{matrix} \text{illustrious and ultimately} \\ \text{triumphant flame in these columns.} \end{matrix} \right] \left[\begin{matrix} A & A & S \\ \text{So, we trust, it} \end{matrix} \right]$
 $\left[\begin{matrix} \text{will prove today} \end{matrix} \right] \left[\begin{matrix} \text{with the campaign for a Permanent November} \end{matrix} \right] -$
 $\left[\begin{matrix} \text{a campaign for reality and for revolution.} \end{matrix} \right]$

(Guardian, 16.3.79).

In this example, the postmodification of the first nominal group ('for a Permanent November') is replaced by the postmodification of the repeated head (campaign) as 'for reality and revolution'. It is notable that the latter case of replacement is not appositional in form: the replacement of the postmodification is a subjective assessment of the head campaign.

In example (30) below the replacement is of the subject and complement of the preceding main clause:

- (30) $\left[\begin{array}{c} A \\ \left[\text{For once in education,} \right] \end{array} \right] \begin{array}{c} S \quad V \quad C \\ \text{this is} \left[\text{an argument where all parties} \right. \\ \left. \text{undoubtedly have only the interest of the children at heart} \right] - \\ \left. \begin{array}{c} S \quad V \quad C \\ \text{it is not sectional.} \end{array} \right]$

(The Weekly Educational Review, 29.3.79).

The lexical function of the interpolation in this example is to deny the suggestion that the parties' interests are sectional. The interpolation therefore postmodifies the whole of the preceding clause.

The interpolation in example (31) consists of a present participle clause which postmodifies the subject of the preceding main clause:

- (31) $\begin{array}{c} S \quad \quad \quad Vi \quad A \quad \quad \quad Vii \quad \quad \quad A \\ \left[\text{Psychologists have too often swung to} \left[\text{the other extreme} \right] - \right. \\ \left. \begin{array}{c} V \quad \quad \quad O \\ \left[\text{sacrificing} \left[\text{the richness and complexity of material} \left[\text{at the altar} \right. \right. \right. \\ \left. \left. \left. \text{of rigour.} \right] \right] \right] \end{array} \right]$

D 305, Block 8, 14-15, p.86).

The interpolation here refers to the consequence of psychologists swinging to the other extreme. That is, it supplies the result of the psychologists' actions, described in the preceding main clause. In fact, the whole participle clause functions as an adjunct of result to the preceding independent clause, which constitutes its evaluation.

The sentence-final interpolation in example (32) contrasts with examples (28) to (31) in that it occurs after the full-stop of the sentence which it postmodifies:

- (32) $\begin{array}{c} A \quad \quad \quad S \quad \quad \quad M \quad \quad \quad H \quad \quad \quad A \\ \left[\left[\text{As will become evident,} \right] \left[\text{this methodological} \left[\text{approach} \right] \left[\text{as much} \right. \right. \right. \\ \left. \left. \left. \begin{array}{c} Vi \quad Vii \\ \text{as any other,} \end{array} \right] \right] \begin{array}{c} Agent \\ \text{is influenced} \end{array} \left[\begin{array}{c} Agent \\ \text{by a particular theoretical position,} \end{array} \right] \right. \\ \left. \left[\begin{array}{c} Agent \\ \text{by a set of assumptions about the relationship between man and the} \right. \right. \\ \left. \left. \text{world, between his consciousness and his activity.} \right] \right] \end{array} \right] \begin{array}{c} M \quad \quad \quad H \\ \text{(One might} \\ \text{only claim that this} \left[\text{approach} \right] \text{has a certain self-consciousness} \end{array}$

about this link between theory and method, /which/ gives it
some slight advantage over those other more naive approaches
/which/ continue to ignore the ways in which methodological
techniques contaminate , and are contaminated by, the very
matters /which/ they seek to comprehend.)

(D 305, Block 9, 16-17, p.7).

Although the interpolation in this example is typographically separated from the sentence it modifies, it nevertheless repeats the main head approach of the subject of the first sentence. The next thing we notice about the interpolated sentence is the repetition semantics of relativisation, represented by three occurrences of the relative pronoun which (shown in the oblique lines). This interpolation is all the more interesting in that it embeds another, smaller, interpolation. The replacement which occurs in this second interpolation is between the two verbs contaminate (active voice), a change from the active to the passive form. This latter interpolation, typographically signalled by correlative commas, comments on the verb contaminate by adding more information in the form of the passive voice. The interruption by the embedded interpolation is of the transitivity relation between the (transitive) verb contaminate and its object 'the very matters which they seek to comprehend'. Although the superordinate interpolation does not delay the sentence boundary (as do examples (28) to (31), it could interrupt the relation between the sentence which precedes it and the sentence(s) which follow it. Fortunately, this long interpolation occurs at the end of its (orthographic) paragraph and so avoids this danger. However, if it were to occur in the middle of a paragraph, then it would create a serious interruption to the semantics of the adjoining clause structure. We will discuss this latter type of complexity in chapter four in the context of clause relations.

Finally, we may consider an example of interpolation which, from the point of view of sentence clarity, seems excessive in its delaying of important parts of clause structure. Thus consider example (33):

(33) (1) There is a lesson here, for the Conservatives especially.

M

H

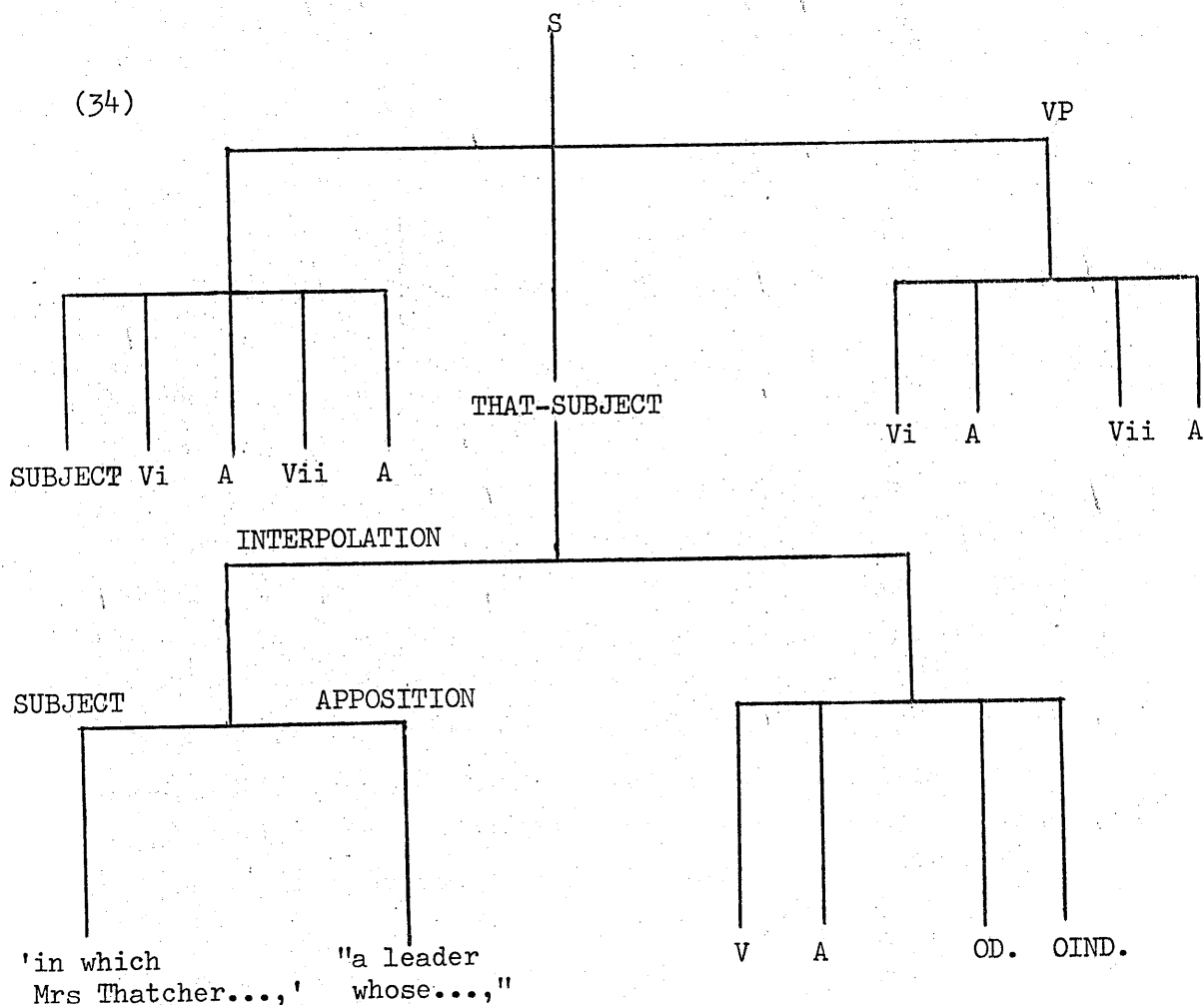
(2) They will not be too much fooled by their own propaganda -

or perhaps more accurately, by the ^M ^H propaganda
^Q which is being put out on their behalf. (3) ^S [Yesterday's
^{Vi} ^A ^{Vii} ^{Viii} ^A radio audience] may well have sensed [from this debate] that
^S [the picture of the present state of politics which is being
^H ^M ^H peddled in some newspapers] - in which Mrs Thatcher, a leader
^Q whose attributes place her somewhere in a line between Demosthenes
^{Vi} ^A ^{Vii} ^{Od.} ^{Oind.} and Queen Boadicea, has already inflicted [mortal wounds] [on a
Labour Government whose only remaining function is to creep away
quietly into a corner and ignominiously die] ^{Vi} ^A is not fully
^{Vii} ^A sustained by the facts.]]

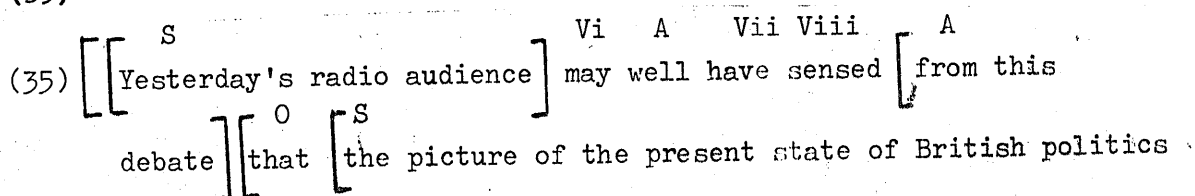
(Guardian, 29.3.79).

The first thing we notice about this example is the sentence-final apposition which occurs between two nominal groups in sentence (2). More specifically, the apposition relation occurs between the two noun heads propaganda. The replacement in the second appositive consists of the noun head propaganda and its postmodification ('which is being put out on their behalf'). The meaning of the relation is explicitly signalled by the adverbial 'or perhaps more accurately.' It is sentence (3), however, which is most interesting from the point of view of syntactic complexity. In this sentence, for example, there is a simultaneous occurrence of apposition and interpolation. In fact, the apposition is embedded within the interpolation itself. The interpolation is explicitly signalled by correlative dashes and interrupts the relation between the subject of the embedded sentences of the matrix-clause ('the picture of the present state of English politics which is now being peddled in some newspapers') and its main verb is. The interpolation itself is a non-restrictive relative clause (signalled by which). An apposition relation occurs between the proper noun Mrs Thatcher and the noun head leader. The second appositive replaces Mrs Thatcher, together with its postmodification ('whose attributes place her somewhere between Demosthenes and Queen Boadicea'). This apposition interrupts the relation between the subject Mrs Thatcher and the verb phrase 'has already inflicted'. In sum, we have here two instances of subject-verb interruption in the same sentence (sentence (3)); in fact, we have a superordinate interruption of subject and verb which subsumes the interruption

by the apposition relation. This idea of a double interruption is perhaps made more clear by the following tree:



A configuration of structures in the sentence such as that shown in example (34) need not necessarily result in a difficult sentence. However, what makes this sentence more difficult than anything we would normally encounter is the sheer length of the interpolation it contains. The main verb is is delayed too long by the intervening relative clause and appositional structure. Because interpolation is, strictly speaking, a non-essential extraneous item of clause structure, we could remove the interpolation in sentence (3) of example (33) to produce example (35):



which is now being peddled in some newspapers ^{Vi} is not fully ^A
sustained ^{Vii} ^A [by the facts.]]]

The main point about example (35) is that we have here preserved the minimal clause structure of the original; consequently, the sentence is still perfectly acceptable to the reader. It does not, of course, contain the detailed specificity of the original, through the loss of the interpolation. However, in the interests of sentence clarity, example (34) illustrates an instance of interpolation which is best reduced or deleted (as in example (35)), even at the expense of some of the sentence's meaning.

In conclusion, the examples of interpolation that we have considered in this section illustrate four important aspects of the role of interpolation within the sentence. Firstly, we have seen that an interpolation may interrupt clauses or the relation between major constituents of the clause; alternatively, it may interrupt the structure of the constituents themselves. In addition, the interpolation may occur at the end of the clause, where it delays the sentence boundary and thus interrupts the relation between the sentence and its adjoining sentence(s). In terms of its mobility within the sentence, interpolation may be seen as a special case of the adjunct function within the clause. This identity with the ordinary adjunct gives rise to the need to develop a new category which is not directly integrated within the structure of the clause. It might well be possible to describe this new category as that of evaluative adjunct (Winter, forthcoming). Secondly, we have seen that the semantic function of interpolation is to evaluate, cataphorically or anaphorically, some element of the adjoining clause structure. This relation of evaluation may often entail replacement meaning between the interpolation and the item evaluated. The latter, however, is not a defining characteristic of interpolation; evaluation is a generic relation that subsumes many modes of expressing authorial comment. This may include a denial or affirmation of the truth of the facts (or situation) evaluated. On the other hand, this evaluation may consist of the addition of new information, in the form of a comment, which in some way judges or assesses some item of the adjoining clause. All such aspects of evaluation have in common a subjective judgement of some kind on the part of the writer; this expresses the essential modifying function of interpolation. Thirdly, it has been shown that a variety of forms of punctuation may signal an interpolation. With the exception of sentence-final interpolations, the punctuation used is correlative in nature, with identical punctuation used

on each side of the interpolation. Finally, we have seen that the grammatical form of an interpolation may range from a nominal group to a sequence of (co-ordinated) clauses, or even more. As example (34) showed, the limit on the use of interpolation depends on the nature and extent of the interruption created by the interpolation. Once more, the latter has to do with an intuitive rule of acceptability to the reader. Interpolation thus has much in common with apposition (see chapter two, section 2.2) but is not limited like the latter in its scope of backward reference.

SUMMARY

In this chapter we have presented the notion of interpolation as an essentially modifying, extraneous grammatical structure. Interpolation is not directly integrated within the structure of the clause and enjoys an adjunct-like mobility within the clause. Like apposition, interpolation interrupts the relation between major constituents of the clause and also the structure of the constituents themselves. The effect of the interruption by interpolation is to delay important parts of clause structure. It is this interruptive role of interpolation which provides scope for syntactic complexity in text. Scope for difficulty in text may also be provided by interpolation at the level of the clause. In the case of the subordinate clause, it is suggested that this type of structure may entail a greater level of perceptual processing because of its contextual dependence. When the subordinate clause occurs as an interpolation, therefore, the difficulty of the sentence is increased by the greater load placed on short-term memory. In contrast, the independent clause, as an interpolation, may introduce a greater level of difficulty into the sentence because of its relative independence of context. For example, the likelihood for difficulty in the sentence is increased when an independent (declarative) clause interrupts the structure of another independent (declarative) clause. The reason for this is that the contextual function of the independent clause is to present new information; when it occurs as an interpolation in another independent clause it creates a **conflict** between the two clauses over the presentation of new information. In this way, it can over-load the information-structure of the sentence.

The contextual distinction between the subordinate clause and the independent clause in terms of the status of the information they present (between theme and rheme) has important implications for the transformational theory of language. T.G.'s analysis of relativisation conflates the non-restrictive relative clause and independent clause in terms of this contextual meaning. We have shown, however, that the two are quite distinct in their contextual roles within the sentence. It is suggested, therefore, that T.G. suffers from an inability to give proper recognition to these quite crucial contextual functions of the clause.

Chapter Four: clause relations

Chapters one, two and three have served to introduce the four principal syntactic categories for analysing syntactic complexity in text. To recapitulate, these categories are: the long subject construction; the nominal group; the appositional group; and, finally, interpolation. Each of these basic categories was postulated as a possible method of approach to the syntactic complexity of individual sentences. The notion of 'individual sentence' is crucial here. The sentence is taken to be grammatically co-extensive with the independent clause and constituted by the following categories:

(1) Subject Verb (Object) (complement) (Adjunct)

The brackets in example (1) illustrate the optional nature of the items they enclose, while the unbracketed items constitute the minimal form of clause structure. The characterisation of the clause in example (1) demarks the context for the complexity that we have so far considered. This discussion has therefore been confined to the boundaries of the single sentence. The potential for complexity of the long subject construction arose from its delaying effect on the main verb of the clause. Excessive modification within the nominal group obscures the relation between heads in the sentence and, also, delays adjoining elements of structure. The potential for complexity of apposition and interpolation, however, arises mainly from their interruption of the relation between major constituents of the clause and also their interruption of the structure of the constituents themselves. This represents one possible 'dimension' of complexity, a dimension delimited by the individual sentence.

A second possible dimension of complexity, however, concerns the relations between sentences in text. This dimension of complexity belongs to the larger discourse function of language, and unlike the first dimension, it has no upper limit to its constituents. Thus, although we are concerned here with interruption between a minimal clause structure of at least two clauses, there is no strict limit to the number of clauses which can constitute this dimension. The study of this type of complexity is therefore primarily concerned with the interruption by apposition and interpolation of structures larger than the individual sentence. Specifically we are concerned here with the structural interruption of clause relations.

Clause relations represent the macro-end of language and their precise characterisation will take up most of this chapter. An important caveat here is that any attempt to define clause relations must begin with a definition of the clause, which lies at the basis of these larger discourse structures. We already possess a definition of the independent clause in example (1) and the subordinate (dependent) clause can be defined in terms of this. A dependent clause may be defined as a clause which comprises the constituents of example (1) but which is signalled by a subordinating conjunction of some kind. Such subordinate structures are unable to stand alone as sentences but must be completed as a minimal context by the independent clause. In sum, the minimum form of a clause relation consists of a sequence of at least two clauses. This may be either a sequence of two independent clauses or a sequence of a subordinate and independent clause. We are now in a position to begin our discussion of clause relations.

The motivation for the search for structures in text beyond the individual sentence stems from a concern to answer the following questions. What makes for the integrity of text and distinguishes it from a random, unordered collection of sentences? What linguistic devices enable us to understand sentences in sequence? This branch of text analysis has to do with the study of cohesion in text. Traditionally, cohesion in text has been studied with reference to such anaphoric/cataphoric devices as pronouns and pro-verbs, etc. These devices constitute finite sets of connective vocabulary and can have nominal groups, predications and clauses etc., as their referents. Because of their finiteness, these sets of connective items comprise an essentially closed-system vocabulary of connection for sentences in text. Winter (1977) organises such closed-system connective items into two principal groups, which he calls vocabulary one and vocabulary two. The vocabulary one items comprise the subordinators and vocabulary two the various sentence connectors.

List (2) below reproduces the Vocabulary One items from Winter (1977, p.14):

Vocabulary 1: the subordinators of English

- (2) After, (al)though, (as though), apart from-ing, as (3), as far as, as well as-ing, at the same time as, on the basis that, because, before, besides-ing, by the time that, by-ing, except that, far-

from-ing, for, from the moment that, given that, granted that, on the grounds that, however, if, (as if), (even if), in addition to-ing, in order to/that, in spite of-ing, in case, instead of-ing, in as much as, no matter how, now that, once, on condition that, provided that, rather than-ing, seeing that, short of-ing, since (2), so that (2), so.... that, such that, so much so that, supposing that, than, that (2), unless, until, what, whatever, when, whenever, whereas, which, while (2), who, why, with the result that, etc.

Correlative pairs: (just) as X so (too) Y, not so much X as Y, not X let alone Y, the er...the er.

The function of vocabulary one items is to connect clauses or to embed one clause within another. In example (2), the numbers in brackets show how many meanings the particular item has in its connective function, while the other brackets indicate that the item in question has more than one meaning. In vocabulary one, the dash followed by ing denotes subordination where the present participle is functioning as the complement of a preposition or as adjuncts in the clause. The subordinators in example (2) are thus arranged into two groups, the first of which is the subordinators of clauses and the second, the correlative subordinators (which connect clause with clause).

The second set of closed-system items, the sentence connectors, is shown in List (3) below:

(3) Vocabulary 2: the sentence connectors of English

Accordingly, in addition, all the same, also, alternatively, anyway, as such, as a result, at any rate, at least, at the same time, basically, besides, in that case, in any case, in such circumstances, in comparison, consequently, in contrast, on the contrary, conversely, correspondingly, differently, equally, essentially, in the event, for example, for instance, for this reason, for this purpose, furthermore, generally, in general, hence, here, hitherto, however, indeed, in effect, in fact, in reply, in return, in short, in turn, in this way, in other words, in spite of this, instead, likewise, meanwhile, moreover, nevertheless, otherwise, on the other hand, in particular, rather, similarly, so (4), more specifically, still, then (3), therefore, thereafter, thereby, there, therein, though,

thus (2), that is, that is to say, to be more precise, up to now, what is more, yet, etc.

Correlatives: not only (but) (also), for one thing for another, in the first place in the second, on the one hand on the other, firstly, secondly, finally, etc.

(Winter, 1977, pp. 16).

The sentence connectors in the first part of example (3) comprise both sentence adjuncts and anaphoric adjuncts. The second part of example (3) shows the correlative pairs of sentence connectors. Unlike the correlatives in vocabulary one, the correlatives in vocabulary two do not necessarily have to have their second member. The second member of the pair may, for example, be omitted in the interests of rhetorical choice.

The closed-system items of vocabularies one and two, the subordinators and sentence connectors, are traditionally thought to contrast with such open-system words as nouns, verbs, adjectives and adverbs. However, Winter (1977) maintains that there is a whole class of open-system words, which he calls vocabulary three, which have similar semantic properties to closed-system words in sentence-connection. The items of vocabulary three are reproduced in List (4) below:

Vocabulary 3: Proposed lexical items of connection

- (4) Achieve, addition, (action), affirm, alike, analogous, antithesis, (attitude), attribute (2), basis, case, cause, characteristic, change, common, compare, compatible, concede, conclude, condition, confirm, connect, consequence, constant, contradict, contrast, conversely, correct, correspond, deduction, deny, depend, differ, differentiate, distinction, distinguish, (do), effect, equal, error, (evaluation), (event), exemplify, exception, (expect), explanation, fact, feature, follow (2), form, function, general, grounds, (happen), hypothetical, identify, instance, instrumental, justification, kind, lead to, like (ness), manner, match, matter (2), mean (2), means of, method, (move), name, (observation), object, opposite, parallel, particular, point (2), (problem), real, reason, reciprocate, repeat, replace, reply, requirement, resemble, respect, result, reverse, same, similar, situation, sort, (solution), specify, state, subsequent, (surprising),

synonymous, technique, (thing), time, truth, unique, way, etc.

(Page 20).

In example (4) the morphological form of the lexical items does not represent the only form they may assume in sentences. For the lexical item compare, for instance, we may expect to find in text such different morphological forms as compares, comparing, compared, comparison, comparative, etc. Example (4) therefore supplies a list of only the stems of the lexical items of vocabulary three, which may be morphologically changed in text.

The lexical items of vocabulary three approximate more to closed-system than open-system words in sentence connection. The basis for this assertion is that these lexical items can paraphrase the connective properties of the traditionally-understood closed-system items of vocabularies one and two. Together, all three vocabularies constitute a special set of contextual items which function as 'signposts' of what a sentence means in connection with its adjoining sentences. More precisely, the explicit occurrence of these items in text signals a particular type of relation between the clause in which it occurs and the clause or clauses which adjoin it. In the case of vocabulary three, there are two ways in which these lexical items may signal a clause relation. Firstly, the lexical items may precede the clause relation, in which case they are said to anticipate the following clause or clauses. Secondly, they may follow the clause relation, in which case there is a retrospective relationship between the clause relation and the lexical item. In both cases, however, the clause or clauses which are so signalled are said to be lexical realisations of the vocabulary item. The lexical items of vocabulary three may have zero-realisation in text. Nevertheless, although in such cases the particular relation between clauses has no explicit indicator, the nature of the relation is 'understood' by the reader. Thus, even though the lexical item compare may not occur explicitly in the clause, it will be nonetheless clear to the reader that the lexical content of the adjoining clauses is being compared in some way. In this sense, vocabulary three items may be called the meta-vocabulary of English insofar as they 'name' the underlying connective semantics of a sequence of adjoining clauses, whilst not necessarily occurring themselves in the clause. The open-system nature of vocabulary three is revealed in their grammatical behaviour as lexical items in the clause. For example, these items may modify or be modified by other lexical items e.

the item example as noun head is premodified by the adjective striking as in: The Yorkshire miners provide a striking example (Winter 77:24) In this function, they take on open-system lexical meaning. On the other hand, the closed-system nature of vocabulary three items is shown in their connective behaviour in discourse.

We are now in a position in our discussion to define a clause relation. A clause relation refers primarily to a binary relation between members. Each member may consist of one or more sentences, where the sentence is co-terminous with the independent clause. The term member thus represents one of a two-part membership rather than a sentence in a one-to-one relation with another sentence. The theory of clause relations is concerned with the kinds of relations between such a two-part membership which can be made explicit by the items of the three vocabularies which we have referred to. The two members of a clause relation together constitute the paragraph, which may or may not coincide with the paragraph orthographically defined. The presence of both members is necessary for an adequate understanding of a clause relation. This means that a sentence which is taken out of context has thereby lost its membership as a clause relation. Winter proposes the latter as a reason for the "intuitive feeling we often have that a sentence has been quoted out of context" (Winter (80) Forum Linguisticum article forthcoming).

This notion of linguistic context is central to Winter's theory of clause relations. Recognition of the importance of studying language in context is not a new phenomenon. Malinowski (1935), for example, was one of the first to recognise the importance of context in language study in his celebrated phrase "context of situation". This was, however, a very broad (physicalist) notion of context; more specifically, it was a behaviourist notion of context. First (1957) widened the scope of Malinowski's notion of context to include the notion of linguistic context. The meaning of Firth's notion of context is thus an amalgam which includes the situational context of an utterance and linguistic context itself. Linguistic context is the result of Firth's poly-systemic view of language as organised into a series of mutually interacting levels: the phonetic, lexical, morphological, syntactic, and semantic. This behaviourist and linguistic definition of context also informs the work of Halliday (1970), which puts forward an essentially functionalist theory of language structure. The latter comprises three functions: the ideational, the textual and interpersonal. Halliday's theory of language structure is also poly-systemic. Within his theory of clause relations, Winter is primarily concerned with linguistic context

which comprises the lexical, morphological, semantic and syntactic components of language structure. The notion of a clause relation is thus a lexico-grammatical concept. Insofar as it describes the contextual relations of the clause beyond the sentence, Winter's theory may be said to have widened even further the scope of linguistic context.

There are two principal rule-governed ways in which we interpret the meaning of one sentence in the light of another. The first is where we match things, actions or people and is called the Matching relation. The second is where we observe a change in time/space sequence. The latter is called the Logical Sequence relation. The great majority of the items of vocabulary one, two and three are included in one of these two relations (or, in some cases, both). In the case of the lexical items of vocabulary three, we can indicate the inclusive nature of the two relations by providing the following (incomplete) list:

MATCHING

Addition
Affirm
Alike
Analogous
Change
Compare
Constant
Contrast
Differ
Fact
etc.....

LOGICAL SEQUENCE

Achieve
Basis
Cause
Connection
Condition
Deduction
Follow
Grounds
Lead to
Means of
Method
etc.....

As an example of the clause relation of Matching, consider the following sentence (given in Winter, 1977 p.7):

(5) They are a contrasting pair of "goats."

This de-contextualised sentence presents a comparison between Mr. Cousins and Mr. Powell in terms of their abilities for political leadership. According to Winter (1977), the main point about this sentence is that although it is grammatically well-formed (it has a subject, verb and complement), it is nevertheless inadequate as information. This inadequacy can be seen all the more readily if we follow the sentence with the question 'In what respects are they a contrasting pair of "goats"?'.¹ In other words,

we do not know that the comparison is in respect of the two men's abilities for political leadership until the sentence is completed contextually by the sentences which adjoin it. The presence in example (5) of the abstract lexical item contrasting (vocabulary three) predicts that the particulars given by the second member of the clause relation will be contrastive. This is clearly seen if we supply the preceding sentence and the two sentences which follow the second sentence of example (5):

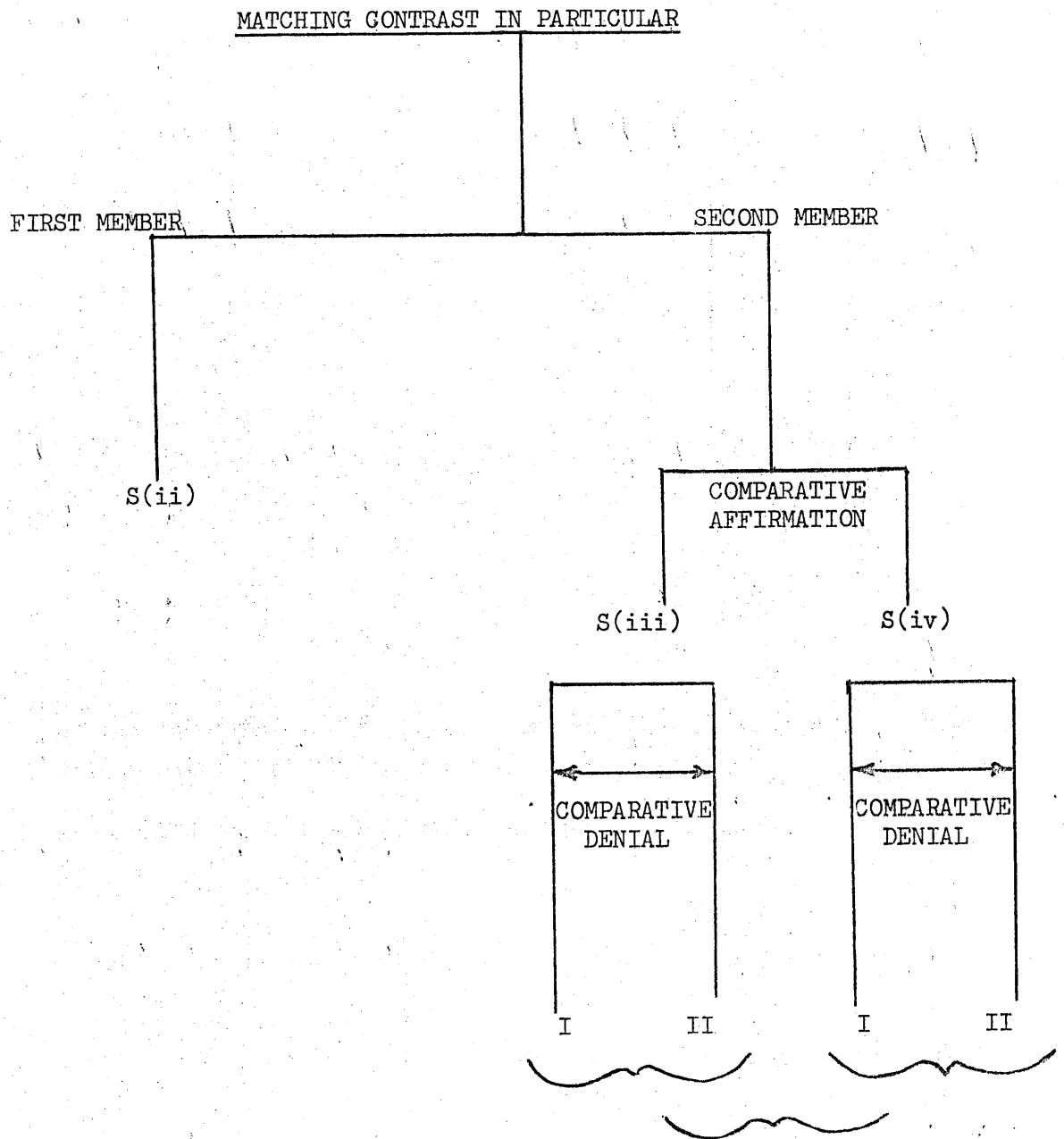
- (6) (i) Mr. Powell is the Tory scapegoat labour has chosen in return for the Tory's choice of Mr. Cousins in the Wilson Cabinet. (ii) They are a contrasting pair of 'goats'. (iii) Mr. Cousins has the big battalions of the Transport and General Workers' Union behind him: Mr. Powell leads a ghost army of Tory philosophers. (iv) Mr. Cousins is very well aware that his troops are men of flesh and blood: Mr. Powell hopes that his ideas will fire the Tory-minded to muster on a philosophical plane.

(Winter, 1977, p. 7) Observer article.

In example (6) sentences (iii) and (iv) supply the details of what is contrasting about the two men, Mr. Cousins and Mr. Powell. The first member of the clause relation is the single independent clause of sentence (ii). Sentence (ii) is said to anticipate the second member of the relation, which is provided by sentences (iii) and (iv). Sentences (iii) and (iv), in turn, contain their own two-part membership of contrast, represented by the two sets of independent clauses connected by the colon. Taken together, all these statements represent what Winter calls Matching Contrast in Particular. The membership of this clause relation is perhaps be made more clear by the tree on page 122.

Matching Contrast in Particular is thus a superordinate clause relation. As example (7) shows, this general clause relation subsumes the relation of comparative denial which obtains between the two sets of independent clauses in sentence (iii) and (iv). That is, the second independent clause in the clause-pair in sentences (iii) and (iv) 'denies' the truth of the assertion presented by the first independent clause. The larger Matching relation also subsumes the relation of comparative affirmation which obtains between sentence (iii) and sentence (iv). Sentence (iv), for example, affirms the truth of sentence (iii). Perhaps the most notable aspect of example (7) is that it illustrates the inclusive nature of clause relations:

(7)



a general relation of Matching or Logical sequence may enclose nested sub-relations within the general framework it provides.

As an illustration of the Logical Sequence Relation, consider the following examples:

- (8) [After the police raids,] [the rifle clubs banned the use of automatic and semi-automatic weapons.]

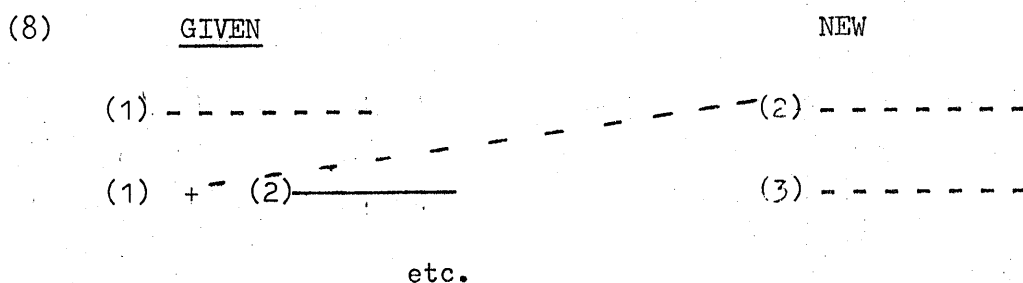
(BBC Radio Commentator).

- (9) [The police raided the rifle clubs.] [Thereafter, the rifle clubs banned the use of automatic and semi-automatic weapons.]
- (10) [The rifle clubs have banned the use of automatic and semi-automatic weapons.] [The move follows the police raids.]

Examples (8) to (10) are taken from Winter (1977, p. 10). Each member of the clause relation of Logical Sequence in these examples consists of one clause (shown by the square brackets), either dependent (as in the first member of example (8)) or independent (as in the first member of example (9)). The vocabulary item which makes explicit the clause relation is different in each example (underlined in each case). In example (8), for instance, it is after, a subordinating conjunction from vocabulary one; in example (9), the connective item is thereafter, a sentence connector from vocabulary two; finally, example (10) contains the lexical item follows (a verb) from vocabulary three. The relation of Logical Sequence in all these examples is so called because the examples are concerned with a deductive time-sequence, namely, that a certain event 'X' succeeded, as a consequence in time, a certain event 'Y'. The striking feature of all three examples is that although the order and nature of the clauses is different in each case, the clause relation itself remains constant. That is, the underlying semantics of all three examples is the same. Examples (8) to (10) illustrate what Winter calls the paraphrase relation between the three vocabulary items. From a contextual point of view of meaning, examples (8) to (10) are clearly different. This contextual difference in meaning can be seen if we use Winter's notion of question-criteria. Here, we formulate a question to which the second member of each clause-pair is the appropriate reply. Take, for instance, example (8). The question which would elicit the second member of this relation as a reply is: 'What did the rifle clubs do after the police raids?' In contrast, the question to elicit the second member of example (10) is: 'Why have the rifle clubs banned the use of automatic and semi-automatic weapons?' The function of WH-questions of this form is to elicit new information. To do this, in its complete form, the question embeds the theme (the information which is already understood), while the WH-pronoun itself corresponds to the form of information requested: an action in the case of example (8) and a reason in the case of example (10). Given this contextual difference between the three examples, however, their propositional content is the same, which makes for a Logical Sequence relation. It is

in this sense that Winter's notion of paraphrase between the three vocabulary items is to be understood, namely, that different grammatical functions (represented by the three vocabulary items) may share the same clause-relating function. Clause relations therefore represent a psycho-semantic process; indeed, as Winter suggests "A clause relation is the cognitive process whereby we interpret the meaning of a sentence or group of sentences in the light of its adjoining sentence or group of sentences." (Winter, 1971, 1974).

In chapters two and three we considered the notion of replacement in relation to apposition and interpolation. Replacement, however, is a much more fundamental concept than this treatment might suggest. It lies at the very basis of the clause-relating function itself and is partly responsible for the cohesion in text. The theory of replacement (see Winter, 1974, and Winter, forthcoming) propounds that the meanings of individual sentences are in a semantic relation with the meanings of the surrounding sentences. The sentences which enter into this semantic relation with a particular sentence are said to form its adjoining linguistic context (Winter, *ibid.*). In any adequate contextual treatment of language in texts, the meanings of adjoining sentences should not therefore be isolated from each other. The important implication here is that part of the meaning of every sentence must be located in the sentences which precede and follow it. The distinction between theme and rheme (between given and new information) in the clause is crucial here.² It will be recalled from our previous discussion of theme and rheme that two important grammatical vehicles are provided by the subordinate and independent clause, respectively. Keen (1978) has also noted that cohesion in text depends on the systematic alternation of the theme/rheme dichotomy to produce what he calls a chaining effect or cohesive tie. That is, the new information presented by one sentence becomes the given information of the next. This notion is illustrated by the following diagram (where the numerals represent clauses):



(Keen, 1978).

Halliday (1970, 1971) has also pointed to the importance of the theme/rheme distinction for relating sentences in continuous texts. According to Halliday, such cohesive ties are part of the textual function of language, where language is used to refer to itself. This distinction is even proposed (Halliday, 1971) as the basis for a stylistic evaluation of narrative text. In other respects, however, the theme/rheme distinction presents only a very generalised and superficial picture of cohesion in text. Winter's theory of replacement refers to the shared grammatical patterns of adjoining sentences in text (see also Halliday and Hasan, 1976). An important 'tie' in text, for example, is the 'systematic repetition' of not only lexical items but also clause structures.* (Winter 1974, and Winter 1980 forthcoming). As we saw in Chapter two (section 2.2),

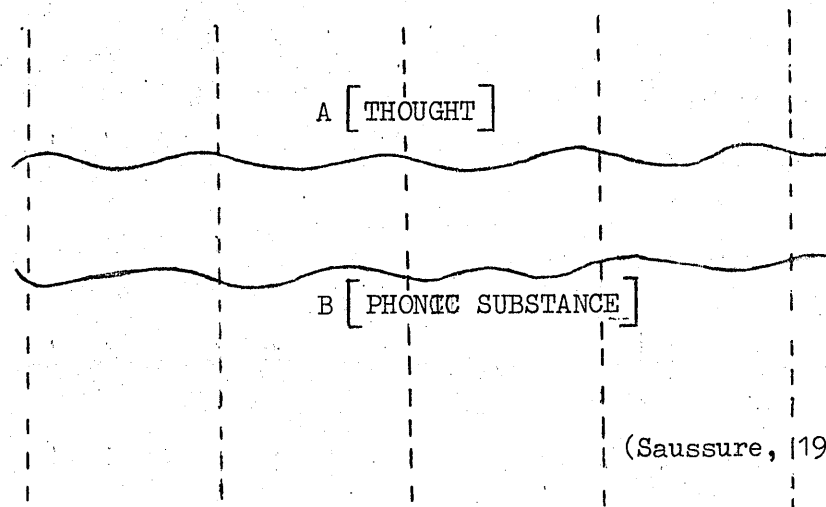
repetition takes the form of deletion, lexical repetition, substitution or a combination of these forms (Winter, *ibid.*). Such repetition of clause structure provides the grammatical framework necessary for the alternation of given and new information in text to operate successfully. The important implication here is that the study of cohesion in text (more specifically, of replacement) must begin with the grammar of the clause. In the parlance of another theory of text analysis (see, for example van Dijk, 1973), any analysis of cohesion must start first at the micro-end of language.

In sum, according to Winter's theory of clause relations, speakers of a natural language such as English share a finite set of structural meanings which enables them to comprehend and produce sentences in sequence. Such structural meanings have explicit realisation in the three vocabulary items already referred to and the repetition semantics of the clause in its replacement function. This closed-system of structural meanings for sentences in sequence acts as a constraint on the number of ways in which we can interpret sentences as a sequenced discourse. The latter is the basis for Winter's assertion that we speak and write in a rule-governed manner.

It might be objected at this point in our discussion, however, that Winter's theory ignores the inherent creativity of language, emphasized so strongly by Wittgenstein (1963) and Chomsky (1965). In Wittgenstein's view, this creativity of language is essentially a semantic creativity, a process of indefinite metaphorical extension of words. For Chomsky, linguistic creativity is a recursive function of a fixed set of grammatical rules, phrase structure rules and transformational rules. Winter's theory of clause relations also allows for creativity, a creativity which inclines

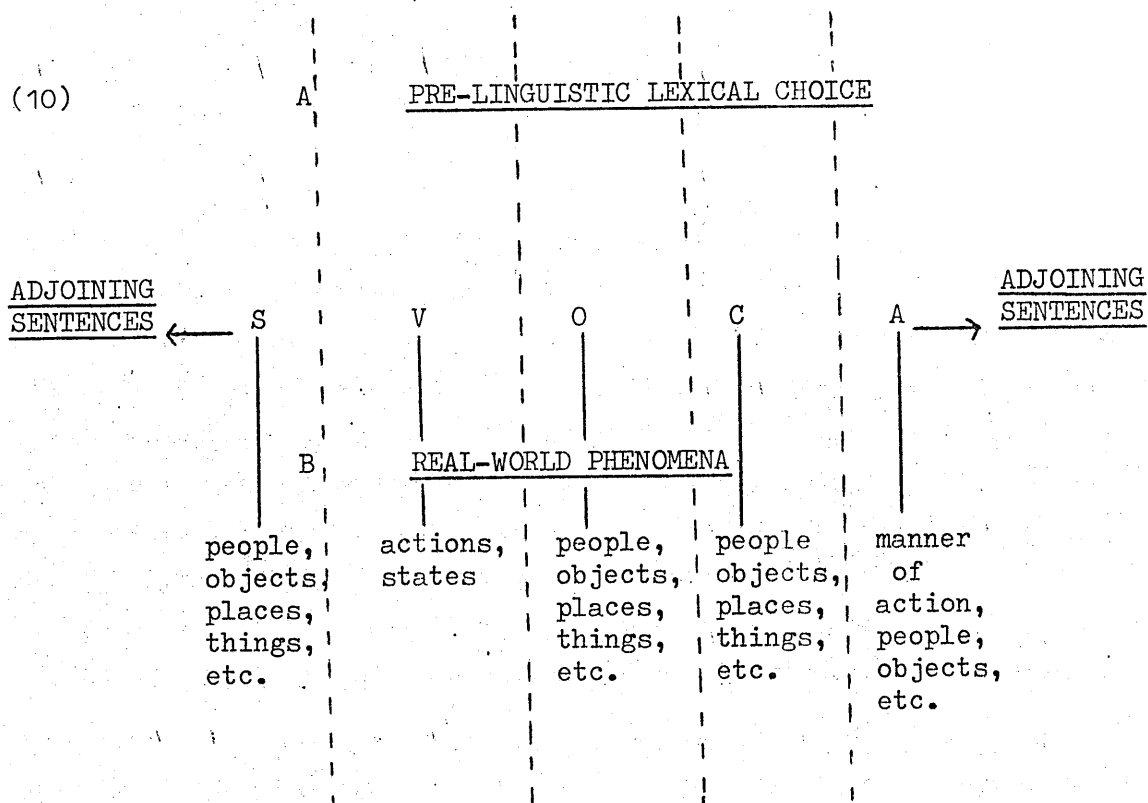
* Systematic Repetition is much more fundamental to the semantics of the clause relation than Halliday's notion of the tie.

more to Wittgenstein's lexical approach than Chomsky's. For example, within the grammatical function of subject, verb, object, complement and adjunct, represented by such open-class categories as nouns, adjectives, adverbs and verbs, there is no limit to the number of things (actions, people, etc.) that we can choose to talk about. Winter contends that there is indefinite scope for lexical choice in this pre-linguistic sense. The linguistic vehicle for our lexical choice is provided by the grammatical context of the clause. The grammatical context of the clause is a formant on the things we wish to say; that is, it makes us talk about things 'in a particular way'. Once in the grammatical context of the clause, we are constrained still further by the clause-relating function of language: the clause inevitably enters into semantic relations with the other clauses which form its adjoining linguistic context. If there were no such semantic relations between adjoining clauses, then we would have no more than a disparate collection of sentences. This is perhaps a linguistically more rigorous statement of the notion relevance referred to by Grice (1968) in his Co-operative Principle.³ In his discussion of the nature of language, Saussure (1974 edition) likens pre-linguistic thought to an amorphous mass; it is, he says, "a vague, uncharted nebula" . (p.112). So, too, for the phonic substance of language. Language, according to Saussure, mediates between thought and phonic substance, giving identity and form to both. Saussure illustrated this idea by the following diagram:



(Saussure, 1974 edition page 112).

The dotted vertical lines in the diagram illustrate the structural delimitation of thought and phonic substance by the mediating role of language. This diagram can be usefully adapted to illustrate Winter's notion of the clause in language. This is shown in example (10) on page 127.



Example (10) is intended to show the mediating function of the clause between the things we wish to talk about (the act of choosing itself, A) and the collection of real-world phenomena from which the choice is made (B). The vertical dotted lines in the diagram denote the structuring role of the clause between these two points of reference. The horizontal arrows are intended to show the syntagmatic relations (ie. clause relations) between the clause and its adjoining linguistic context. Specifically, the lexical form of the clause to some extent determines or constrains the lexical form of the succeeding clause(s). The latter accounts for the strong predictive role of the clause in discourse.

It has been necessary to deal with clause relations at this length in order to supply the necessary background material to our discussion of complexity beyond the clause. It is hoped that the examples which follow will show that complexity beyond the boundary of the clause in large part involves the interruption of Winter's paragraph. That is, there is scope for apposition and interpolation to interrupt the two members of a clause relation, thus delaying the lexical realisation of the second member. There is also scope, of course, for the long subject construction and a succession of long nominal groups to delay the second member of a clause relation.

The possibilities for interruption of a clause relation are perhaps especially striking in the case of apposition and interpolation, since

these are largely extraneous categories to (minimal) clause structure. In this instance, consider example (11) below (see example (22), chapter two, page 109):

- (11) Some years ago Stanley Cohen and I began some research at the maximum security wing of Durham prison. (At the time we were particularly interested in the deviant status of those who occupied the wing; many of them were notorious criminals and our almost unique opportunity to interview them promised to provide some interesting criminological insights into the meaning that their behaviour had had for them.) (2) However, as the months went by (overall we visited the wing for a period of four years), we began to lose interest in the nature of their criminality and became increasingly absorbed by the problems which they faced in coping with their present reality; [not so much^M the evident physical
^H properties ^Q of living in the wing] - ^M the locks, ^H the landings, ^M the
^H guards on every corner, ^Q the oppressive heat, ^M the lack of ventilation
- [but rather the type of inner subjective life which developed
in response to these oppressive circumstances.]

(D 305, Block 9, 16-17, p.13).

The broken line in sentence (2) indicates five nominal groups which are in apposition to the preceding first member of the clause relation of Matching. The latter is shown in square brackets. The second member of the relation, which occurs after the apposition, is also shown in square brackets. The various nominal groups in the second appositive clause replace the nominal group 'the evident physical properties of living in the wing' in the first appositive (indicated as usual by the rectangular box and superscripts). In more precise terms, the clause relation interrupted by the apposition is that of Matching Contrast. The underlying semantics of this relation is that of (partial) error and (partial) correction. That is, the first member of the relation describes what is 'not so much' the case with life in the prison, while the second member describes what is the case. The clause relation itself is made explicit by the two correlative co-ordinators 'not so much' and 'but rather', both of which are subordinators from vocabulary one (underlined in example) (11).

It is certainly not suggested here that example (11) represents a complex sentence. It does, however, suffice to illustrate a potential context for complexity, especially if the appositional structure is longer and more complex. The type of complexity which example (11) (potentially) illustrates concerns the temporary defeat of the reader's expectation of the second member of the clause relation by an extraneous structure (in this case, apposition). The expectation on the part of the reader is developed by the prediction set up by the first member of the clause relation that the second member is about to follow (or be lexically realised). The prediction in example (11) is made explicit to the reader by the co-ordinator 'not so much'. As Winter (1977) suggests, "clause relations is a system of predictability of context; that is, given one sentence with its preceding context, the lexical selection of the next sentence is frequently predictable, and this predictability is the crucial part of the semantics of the clause relation" (page 35). This notion of reader expectation is supported by Leech (1969) with reference to foregrounding in poetry. Thus, in the case of an interruption of a clause relation, the reader has to hold in short-term memory the prediction set up by the first member of the relation while he processes the intervening extraneous structure. In short, the interruption postpones the second member of the relation.

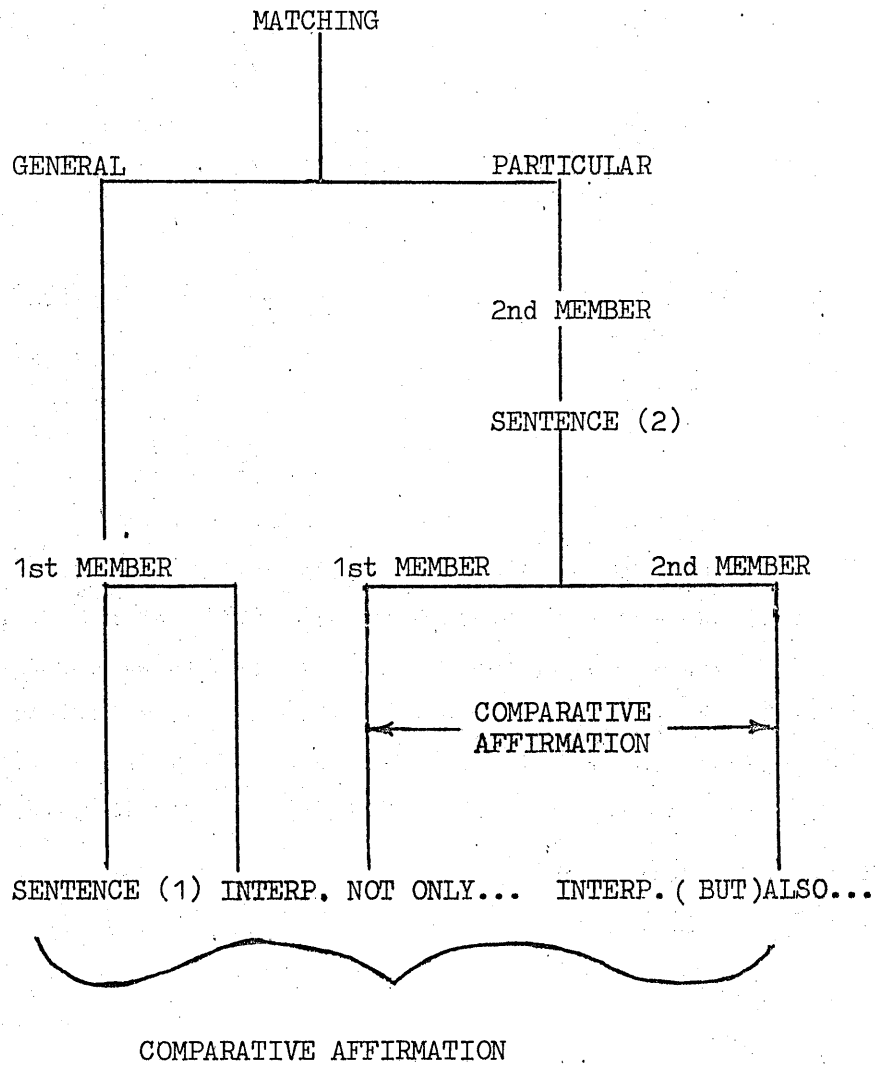
Another example of the interruption of the clause relation of Matching is provided by example (12) below:

- (12) [Ernie was a phenomenon] - one of the last great freelances working in what the French call 'actualite'. (2) [He was not only a great reporting cameraman] - which his awards recognised (they included Cameraman of the Year for his portfolio of Vietnam films with Julian Pettifer), [he was also a remarkable loner, a man who believed in owning and flying his own plane anywhere in search of a visual story.]

The Listener, 15.3.79).

As we have already analysed this example (see the discussion of example (1), chapter three, section 3.1), we can ignore the faulty punctuation here. The underlying semantics of the Matching relation in sentence (2) is that of comparative affirmation. The relation is explicitly signalled by the correlative pair of sentence connectors (vocabulary one) 'not only' and 'also' (both of which are underlined in sentence (2)). Strictly, the fully-explicit form of the second sentence connector is 'but also', as in the structure: 'not only X but also Y'. The deletion in the second member of this first part of the connector is a function of the repetition of the subject 'he' which begins the second member of the relation. As such, the deletion illustrates how the realisation of the explicit connectors of a clause relation is often a matter of rhetorical choice on the part of the writer. There exists a compatibility relation between the first and second members of sentence (2). The second member affirms the truth of the first by supplying compatible particulars in the comparison. The second sentence of example (12) thus illustrates the interruption of a clause relation of Matching by a double instance of interpolation (denoted by the broken line in the example). However, this clause relation in sentence (2) is itself embedded within another, more inclusive, clause relation, namely, that of General and Particular. The independent clause of sentence (1) (shown in square brackets), for example, expresses the general statement that 'Ernie was a phenomenon'. The occurrence in sentence (1) of the abstract noun 'phenomenon' predicts (anticipates) that the following sentence or sentences will provide the particulars for this generalisation. The particulars for this generalisation are lexically realised in sentence (2) by the clauses contained within the square brackets. The first independent clause of sentence (2) predicates of the subject 'he' the quality of being 'a great reporting cameraman' and the truth of the assertion of this quality is then affirmed by the compatible particulars of the clauses of the second member. Taken together, the two members of sentence (2) affirm the truth of the general statement of sentence (1). In example (12), therefore, there is a double occurrence of comparative affirmation and two sets of clause relations, one embedded inside a superordinate relation. This inclusive relationship can be seen more readily in the following tree on page 130, example (13).

(13)



Example (13) also illustrates the evaluative interpolation which occurs at the end of sentence (1). The effects of a sentence-final interpolation of this type are two-fold: firstly, it delays the sentence boundary; and, secondly, it interrupts the relation between General and Particular. Example (12) is interesting, therefore, in that it illustrates two instances of interruption within the same (inclusive) clause relation.

Another example of interruption of the Matching relation is provided, finally, by example (14) below. Just as in example (12), example (14) illustrates comparative affirmation by a comparability relation between the first and second members:

(14) (1) Walton gaol has survived its second emergency in as many months.

(2) [It is not just here that prison officers are angry and

disappointed that Mr. Justice May's inquiry into the prison service has not produced an interim report on pay] - although, so far, it is only here that they have expressed that impatience in industrial action which has kept prisoners locked up for virtually 23 hours out of

24. (3) [The Prison Officer's Association \emptyset lost its bid for a truce to industrial action while the inquiry was sitting; local branches are free to go ahead with action on what they take to be local issues.]

(Guardian, 21.3.79).

The first and second members of this Matching relation are indicated in square brackets. The relation between the two members is interrupted by the sentence-final interpolation in sentence (2). The interpolation itself takes the form of a subordinate concessive structure which comments on the first member of the relation. The second member of the relation affirms the truth of the first by providing compatible particulars. The truth of the assertion that the prison officers are angry is affirmed (matched) by the compatible assertion that the Prison Officers Association is also angry at the deficiencies of Mr. Justice May's inquiry. However, while the first member is explicitly signalled by the sentence connector 'not just' (vocabulary two), the second member is not so explicitly signalled. The fully-explicit form of the relation is 'not just X (but) Y also'. The second member of the pair here has therefore been deleted from the second member of the clause relation in example (14). The latter is indicated in the third sentence of example (14) by the zero symbol \emptyset . The second member is also compatible with the first in terms of its main verb and complement. The second member could, for instance be rewritten: 'The^S Prison Officers Association is also ^Vangry^C....', with everything that follows functioning as a subordinate because-clause of reason. Thus, although the sentence connector in sentence (3) is not explicitly given, the relation is nonetheless clear to the reader because of the prediction set up by the

connective 'not just' in the first member. The lexical realisation of this second member, however, is delayed by the interpolation at the end of the first member.

The purpose of the analysis of examples (11), (12) and (14) is not to try and provide illustrations of syntactic complexity or 'difficult' syntax, but is much more modest than this. It is, rather, to show possible contexts for syntactic complexity. It is suggested here that such possible contexts are provided (in longer discourse structures) by the interruption of clause relations and the consequent delay of the second member by extraneous structures. Interpolation and apposition, as we have seen, are possible candidates for this type of interruption. So, too, are the long subject construction and the complex nominal group. The use of the latter structures, for instance, not only delays important constituents of the clause but can also function to delay the sentence boundary. By delaying the sentence boundary, they may possibly delay a member (or members) of a clause relation. We have seen how clause relations are responsible for much of the cohesion in text and how relations can be nested in larger relations. The excessive interruption of these clause relations, it is argued here, has grave implications for the cohesion of text. Excessive interruptions impairs the system of predictability in text and therefore renders the reader's task more difficult. It does so by systematically impeding the predictive task which is required for successful reading. (Cf. Leech's notion of defeating the predictive task. Leech 1969)

SUMMARY

This chapter has attempted to situate the syntactic categories examined in chapters one, two and three (especially apposition and interpolation) in a larger discourse context. One possible dimension of complexity produced by the long subject, nominal group, apposition and interpolation concerns the interruption and delaying of the structure of the clause. A second dimension of complexity is proposed, however, where the use of such categories in the clause can interrupt and delay the membership of a clause relation. Clause relations concern the grammar and semantics of cohesion in text and supply the majority of the cohesive ties which we rely on for successful reading. Clause relations are also seen to have important implications for the psychological basis of language. The individual clause provides a (grammatical) controlling constraint on our lexical creativity. This controlling constraint does not limit what we say or wish to talk about, but the way we communicate it. A secondary constraint concerns the relationship between clauses, as manifested in clause relations. The latter are responsible for the meanings of sentences in sequence. 'Creative communication' is therefore regulated by the grammar of the clause and the larger clause context. Finally, clause relations have to do with the system of predictability in language structure. A possible cause of complexity in text, it is suggested, results from the impairment of this system of predictability.

Chapter Five: Conclusion

The basic assumption underlying the approach to complexity adopted in this thesis is that there is something more to complexity in text than merely sentence-length or semantic density. If we accept this assumption, then we are faced with the task of 'explaining' this complexity from some particular frame of reference (e.g., the psychological, graphological, etc). The frame of reference chosen here is linguistic; more specifically, it is a grammatical analysis which takes as its primary data the configuration of structures within the clause. However, any grammatical analysis which purports to explain complexity must have at its disposal a rigorously-defined set of categories or descriptive procedures. Chapters one to four are attempts to provide such a necessary set of categories. They also represent an attempt to delimit the scope of the inquiry in its early stages. This limitation of scope is represented by the boundaries of the clause.

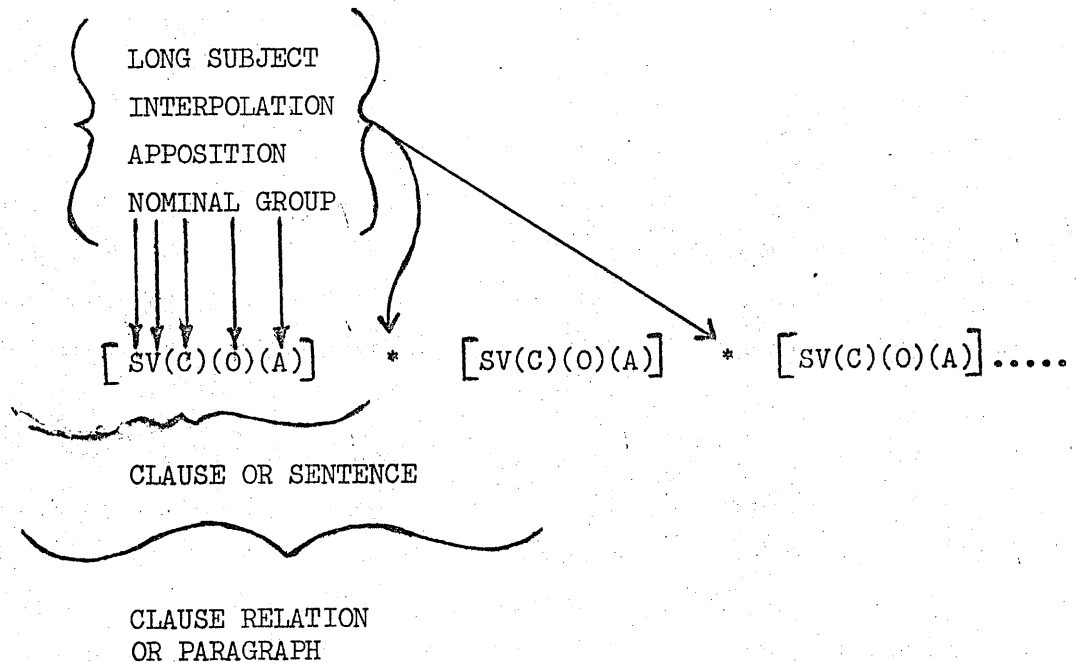
However, what is the empirical basis of such categories of description? There are two possible methods of approach here. Firstly, we can appeal directly to our 'intuition', as proposed by Chomsky (1965), and formulate categories on the basis of our knowledge as native speakers of the language. According to this method of inquiry, the analyst is allowed to produce a set of generalised statements about linguistic structure on the basis of his own introspective judgements. An alternative approach, adopted by sociolinguists (e.g., Trudgill, 1973, Labov, 1966) is to derive categories of description from the empirically-observed use of language in its real-world context. The difference between these two approaches hinges on a distinction between seeing the categories for linguistic analysis as competence-based (as in Chomsky's view) or as performance-based (as proposed in sociolinguistics). The view of linguistic description adopted here is a compromise between these two extremes. For example, the linguistic categories that we have proposed are, initially, the result of intuitive insight and knowledge of the language. However, they cannot in any sense be said to naturally inhere in language, merely waiting to be discovered by the linguist. They are hypothetical constructs which must be refined and tested against the language data. The texts to which such categories are applied represent (ideally) as broad a spectrum of the language under study as possible. The final set of categories therefore results from generalisations made from the data collected. The overall purpose of this approach

is to produce a set of descriptive procedures which can be subjected to psychological or statistical validation.

This is the stage reached in the analysis of the present work. Such categories as apposition and interpolation have been applied to English text; on this basis, they have been refined and generalised as descriptive procedures for the analysis of syntactic complexity. The long subject and nominal group are much easier to define (in theoretical terms) than apposition and interpolation. The former are much more clear-cut grammatical categories than the latter which, as we have seen, entail quite subtle semantic relationships. The long subject and nominal group therefore do not need the same extent of application to texts to be proposed as possible categories for syntactic complexity. An important implication of this is that before we attempt statistical or psychological tests, we must first have well-defined categories to test.

The point of departure for all the categories that have been proposed is the clause. The categories must be applied in relation to the context of the clause and its parts. The structure of the nominal group, for example, must be analysed and assessed in relation to the heads of the other constituent structures of the clause (e.g., S, V, O, C, and A) and to what extent it obscures or complicates the relation between these. In the same way, the long subject must be analysed and understood in relation to the part of the clause which it delays, namely, the main verb or verb phrase. Finally, the categories of apposition and interpolation must be analysed as interruptions to particular parts of clause structure; we must ask what part of clause structure is being evaluated or what item in the clause is further specified by the modification. We must also ask what is the nature of the relationship between these modifiers and the parts of the clause interrupted. In short, the analysis proceeds by investigating the interaction of these categories with the constituents of the clause. As we saw in chapter five, the clause also provides the basis for larger discourse structures, namely, clause relations. In this part of the analysis for complexity, we are concerned to examine how structures beyond the sentence can be interrupted or delayed by the categories that we have considered. Since there is, theoretically, no maximum limit to the number of sentences contained by each member of a clause relation, there is no maximum limit to the size of the structures which can be interrupted. For example, the structures which interrupt a clause relation could themselves consist of a long succession of sentences with long subjects or complex nominal groups, etc. In such cases, the possibility of syntactic complexity

or difficulty in text would be very great indeed. The essentials of what is envisaged in the model of complexity proposed here is summarised in the following diagram:



This diagram does not, of course, show all the possibilities for interruption. In the diagram, the arrows indicate the place of interruption, while the asterisks denote the interruption of clause relations. The diagram therefore illustrates the centrality of the clause as the point of reference for larger discourse structures.

What has been said about syntactic complexity or difficulty in text often being a function of structural interruption (e.g., by apposition and interpolation) is perhaps best understood by reference to the notion of 'noise' in text. Insofar as they are mostly extraneous, modifying structures, apposition and interpolation are, strictly speaking, non-essential elements of minimal structure. The 'difficulty' that the excessive use of such structures gives rise to could be described as 'noise within the grammar of the clause or paragraph.' However, it is an over-simplification to suggest that we should have either a grammatical or a semantic approach to the problem of complexity or difficulty in text. Ideally, we should have an approach which combines both frames of reference. Nor is it tenable to suggest from this that optimum clarity is achieved in text by a familiar and simple syntax or a low level of lexical content. It seems much more reasonable to adopt the view of Goodman and Greene (1977). Goodman and Greene, for instance, suggest that if the content of material in text is

too difficult for readers, they will have difficulty in using the grammatical structures to gain maximum comprehension of the content. On the other hand, if the grammatical structures are too difficult or unfamiliar to readers, they will experience difficulty in translating the structures into language with which they are familiar. The authors suggest that this has the following pedagogical implications. If the aim of the instruction is to introduce new concepts and ideas to the reader, then the syntactic structures should be reasonably clear and uncomplex. The latter removes from the reader the task of processing difficult syntactic structure at the same time he is required to assimilate new ideas. On the other hand, if the instruction is attempting to introduce unfamiliar or more difficult syntactic structures, then the lexical content of the material should be familiar to students.

Inevitably, it will be objected by some that a grammatical approach to complexity of the sort proposed in this thesis has no validity at all.¹ The substance of this type of objection is the argument that the linguistic constructs proposed have no relevance to the processing tasks facing the reader. The same type of objection has been put forward by literary critics (e.g. Hirsch, 1972) against the approach advocated in linguistic stylistics. Specifically, the controversy between the two schools has to do with the question of how to characterise authorial style: is authorial style the result of the lexical content of a work (and its effects on the reader), or is it the result of its syntactic form and the linguistic choices manifested by the writer therein? A standard reply from the province of linguistic stylistics is provided by Pierce (1977, P.56):

The achievement of a successful description of a text is that it formulates intuitive insights into an aspect of meaning of a language event in an objective, coherent, consistent and 'falsifiable' manner in a way which is intuitively satisfying

(my inverted commas)

Further,

This complex integration of observations generalising from specific details and a small set of specific and unique statements of a high degree of generalisation constitutes the description and interpretation of a text, one statement of its meaning. (p. 67).

The word 'falsifiable' in the first quotation is very important as it is the principal justification for characterising authorial style in terms of the 'form' of a work. For example, it is possible to confirm or disconfirm one interpretation of a literary work on the basis of the linguistic details adduced in support of the interpretation. The interpretation is rejected if it is not borne out by the linguistic details of a work. However, it is difficult to see how interpretations evolved on purely intuitive grounds can be tested as true or false. By the same token, it is suggested here that a linguistic characterisation as a partial explanation for complexity and difficulty in reading has the advantage that it is falsifiable. That is, it may be rejected as unsound by alternative examinations of the linguistic details. More importantly, the categories which it proposes may be confirmed or disconfirmed by psychological tests. Clearly, we bring our world knowledge and personal biography to the reading process but this does not explain away the grammatical knowledge that we must also bring to bear in reading. Because we segment the continuous stream of prose and speech in highly specific ways (excepting the use made of typographical and graphological signals), we must, in consequence, have recourse to grammatical categories of some sort.

In conclusion, the linguistic categories that have been proposed in this thesis are offered as tentative apparatus with which to analyse complexity in text. They are proposed in the absence of an adequate set of linguistic categories. The theoretical status of apposition and interpolation is that of inductive generalisations based on the analysis of their behaviour and use in prose texts. All the categories are put forward as candidates for psychological testing. To this end, it is hoped that their confirmation or disconfirmation may enlarge our understanding of the reading process, not just at the sentence level but at a larger discourse level. At the larger discourse level, such an adequate set of categories is perhaps all the more necessary if we accept Pierce's contention that "stylistics suffers from the inability to produce adequate linguistic descriptions of texts as coherent units rather than as strings of sentences held together by an intermittent network of reference and parallelism"(P.6 and P.213)

NOTES TO INTRODUCTION

1. A more extensive discussion of I.C. analysis is provided in Grammar (Pelican, 1971) by Frank Palmer, see pages 124-134.
2. See, also, Chomsky's scathing review of Skinner's Verbal Behaviour, in Language, 35. (1959) pp. 26-58.
3. The account of transformational-generative grammar which follows (i.e. of the generative and interpretive semantics debate) is taken from Leech (1974) and Sampson (1975).
4. For a discussion of these three hypotheses see Formal Grammars in Linguistics and Psycholinguistics (1974, Mouton) Vol. 3, W.J. Levelt.
5. See, also, G. Sampson's review of J.S. Ganz 'Rules' in Lingua, 32 (1973) pp. 160-4.
6. P.F. Strawson, in Introduction to Logical Theory (Methuen, 1967), however, points to the occasional mismatch between a conjunction of propositions in the propositional calculus and conjoined clauses in natural language. This is particularly the case where temporal relations are concerned. For example, consider the following conjunction:

(i) P & Q

The truth-value of this conjunction remains constant if we reverse the order of the propositions as follows:

(ii) Q & P

We may assume that the conjunction in example (i) corresponds to the conjunction of propositions in example (iii):

(iii) John fell ill and died.

If we reverse the order of these propositions (as in example (ii)), this produces the following false conjunction of propositions:

(iv) John died and fell ill.

- (7) Examples (22) and (23) are taken from 'Grammar and Reading in the Classroom' by Yetta Goodman and Jennifer Greene, in Linguistic Theory: what can it say about reading? (The International Reading Association, 1977).

Notes to chapter one

1. The majority of the examples of the long subject given below did not originally display such a marked word-order, especially example (8). For convenience of exposition, their original word-order was changed so that the verb phrase occurred in sentence-final position. However, in the light of sentences such as example (8), this should be taken as further proof of the tendency of English syntax to delay or interrupt excessively-long subject structures.
2. Co-ordination is also used, of course, in the formation of compound nominal structures in subject position.
3. The example that Chomsky gives of a predominantly left-branching structure (1965, p.10) is itself a good illustration of the long subject. The sentence given is:

S
"Quite a few of the students who you met who came from New
York ^V ^C are friends of mine".

4. In the case of written discourse, however, such sentences would be confined more to idiomatic expressions or habitually-encountered structures. In the case of novel sentences, the speaker is much less likely to know how they will unfold.
5. Such suprasegmental properties of language include stress, rhythm and intonation. These features of language are largely responsible for the organisation superimposed in the smaller linguistic segments (such as the phoneme and morpheme.)
6. John Lyons (1970, p.91) notes that the depth hypothesis of Yngve's is "almost certainly incorrect ... since it is not clear that left-branching structures are as difficult for human beings to 'process' as they should be according to the hypothesis." In support of his contention Lyons refers to predominantly left-branching languages such as Japanese and Turkish. Perhaps the best way to answer Lyons' objection is to consider an analogous question in the articulatory production of language. It is well-known, for example, that any (normal) child is capable at birth of learning any attested (human) language. Providing that the

environment is conducive to learning, he can master the articulatory task of producing the vowel and consonant phonemes of the language. This is the case even though the vowel and consonant phonemes between languages (and hence, the articulatory task facing the learner) are manifestly different in many cases. It is equally well-known, however, that mature native speakers of one language experience great difficulty when trying to articulate the sounds of another language whose sound system is quite different. Imagine, for example, the task facing an Englishman trying to produce the sounds of languages like the Xhosa of Africa or Arabic. While not impossible, it is a far from easy task. The reason for this is that although the vocal apparatus is the same for all humans, it becomes habituated to the articulatory form of the sound-system it is exposed to from birth. Thus, while in the early stages of learning the vocal organs are capable of producing any sounds, this facility is lost with maturity. Now, by the same token, the speaker's perceptual mechanism becomes habituated to the grammatical system it is exposed to in the early stages of development. Meaning is apprehended in the way dictated by this grammatical system. Accordingly English speakers are used to structures where the main verb is not delayed too long, while for German speakers this is quite the reverse. While in German left-branching structures are quite usual and relatively unrestricted, in English there is a marked tendency for right-branching structures. Yngve's depth hypothesis therefore seems partly correct when applied to English, even though it is not appropriate for other languages. In consequence, Lyon's criticism seems to be unjustified in this respect.

7. Quirk et. al. (1972) claim that the devices of postponement (what we have called delayed subject) serve two principles: that of end-focus and end-weight. The delayed structures which we have discussed (where we front Anticipatory It and delay the subject (S_2) till after the main verb) is dealt with by Quirk et. al. as extraposition of subordinate nominal clauses. The authors argue that the postponed position is more usual than the orthodox position before the verb (Sec 9.64). This is precisely the point captured in our notion of long subject. The long subject is marked by virtue of its unusualness. While the long subject constructions we have been discussing, however, are marked, the postponed structures given in Quirk et. al. would be less marked by our criteria. Some of the postponed structures cited as examples are given as follows:

S1 Vi Vii S2
(1a) It doesn't matter what you do.

S1 V O S2
(2a) It surprised me to hear him say that.

S1 V O C S2
(3a) It makes her happy to see others enjoying themselves.

The more marked corresponding long subject forms of these examples are as follows:

S Vi Vii
(1b) What you do doesn't matter

S V O
(2b) To hear him say that surprised me

S V O C
(3b) To see others enjoying themselves makes her happy.

In sum, the principle of extraposition of clausal subjects in this way is, once again, a function of clarity of syntax rather than merely end weight. This seems to be recognised to some extent by the authors who admit that "In many cases, the postponed elements no doubt undergo postponement because their length and complexity would otherwise lead to an awkwardly unbalanced sentence." (p. 967).

Notes to chapter two

1. See, for example 'The language of Legal Documents', in Investigating English Style by David Crystal and Derek Davy (Longman, 1969). Crystal and Davy argue that much of the difficulty in reading legal documents arises from the frequent use of long nominal groups. For example, they claim that one oddity of legal language is the habit of inserting postmodifying elements at those points in the nominal group where they will produce the most unambiguous meaning. However, they conclude that "the concern of legal language for logical structure makes a reader's task more difficult, and that of anyone attempting to 'sight read' such passages (as in reading aloud) almost impossible." (p. 206). It is significant that they do not see complexities of this kind as being confined to legal language.
2. What we have been saying about the scope for indefinite repetition of the premodification and postmodification of the noun head is part of the recursive potential of human language, according to Chomsky (1965). The term recursion derives from mathematics but was adapted by Chomsky as part of the system of rules in grammars of natural language. Thus, the premodification of the noun head in a Chomskyan grammar would be assigned the following type of recursive structure at position E (represented here by phrase structure rules):

premodification \longrightarrow (D) (O) ((E) (X)) (N) N.H.
D \longrightarrow determiner/predeterminer
O \longrightarrow ordinal/cardinal numeral
E \longrightarrow adjective
X \longrightarrow E (X)
N \longrightarrow noun
N.H. \longrightarrow noun head

By the same token, the postmodification be assigned a recursive structure as follows:

postmodification \longrightarrow N.H. ((p.d.) (Y))
N.H. \longrightarrow noun head
p.cl. \longrightarrow postmodifying clause
Y \longrightarrow p. cl. (Y)

3. Replacement is a function of the repetition of the clause. (Dixon 1964, Quirk et al 1972, Winter 1974, and 1980 forthcoming Forum Linguisticum). In Winter 1977, Winter defines replacement as: "Replacement is a function of the repetition of the clause; it means offering a new lexical element for the repeated clause so that it maintains its lexical uniqueness with respect to the clause which it is repeating." (P.31).

S V C C

Evangelecologists are good at publicity; bad at getting things done.

Here the replacement is a function of the repeated structure of S V by deletion (Evangelecologists are) and the repeated structure of the prepositional phrase at which postmodifies the adjective heads good and bad respectively. What is replaced is the adjective complement good by bad, and the 'objects' of the preposition at in which the nominal head publicity is replaced by the gerundial nominal getting things done. The antonym bad signals a contrast for the second clause. We could regard the example as an answer to the question: "What are Evangelecologists like at publicity (and) getting things done?"

4. The terms unspecific versus specific declarative clause comes from Winter (forthcoming Allen and Unwin). He says that studies of English grammar will have to take into account this contrast of information status between two clauses. In investigating the way in which meaning is signalled in unspecific declarative clause, we will have to take into account the way in which this clause is lexically realised by its adjoining specific declarative clause.
5. Although such adverbials are used to make explicit the appositional meaning of the relation, their presence in a sentence does not always denote appositional meaning. They may, for example, denote an interpolation instead (see chapter three). Generally, however, their presence indicates apposition.

Notes to chapter three

1. The discussion of Interpolation in this chapter closely follows the description given in Winter (Allen and Unwin forthcoming), which treats the grammatically interruptive nature of interpolation (also called parenthesis) as a phenomenon of great theoretical importance to English Grammar.
2. Although the present work is specifically concerned with the analysis of syntactic complexity in 'written' text, many of the suggestions made and points raised apply equally well to spoken discourse. In this secondary sense, the conclusion made here may in many cases be construed in relation to discourse in general.
3. Interpolation may be seen as a special case of interjection. Like the latter, interpolation enjoys a distributional range which is denied to the major constituents of the clause. For a discussion of interjection see Winter (Allen and Unwin forthcoming).

Notes to chapter four

1. Questions provide the criteria by which to judge the nature of clause relations (Winter, 1974). Questions, for example, form the items to which the second member of the clause relation is a reply. Thus, the paragraph is contextually dependent on the adjoining paragraphs: it is a reply to the question(s) posed by the adjoining sentences.
2. Broadly, the distinction between theme and rheme corresponds to the subject and predicate, respectively. Equivalent terms used for this distinction are Topic and Comment. See, for example, Semantics: a new outline, F.R. Palmer (Cambridge University Press, 1976) for a useful discussion of this distinction. The Prague School, however, have attempted to refine this distinction between theme and rheme. They have developed the notion of Functional Sentence Perspective, which defines theme and rheme in terms of communicative dynamism, or degrees of importance attached to the information in the sentence. For The Prague School, the distinction is not as clear-cut as for some linguists but is a relative concept. According to F.S.P., the theme and rheme of a sentence may vary as a result of the extra-linguistic (situational) context, or as a result of the purpose of the communication, etc. For a discussion of this concept see J. Firbas: 'On defining the Theme in functional Sentence Analysis' in Travaux Linguistiques de Prague, Vol. 1. 267-80.

For van Dijk, the topic of a sentence is defined as part of the underlying representation which is identical with the underlying representation of the preceding sentence. See, for example, T. van Dijk's Some Aspects of Text Grammar: a Study in Theoretical Linguistics and Poetics (The Hague Mouton, 1972). For a transformational-generative approach to the question of topic and comment in the sentence, see also N. Chomsky's 'Deep Structure, Surface Structure and Semantic Interpretation', reprinted in Steinberg and Jakobvits, 1971. For Halliday, however, given and theme are quite separate. Given concerns what you 'were' talking about, while theme means what you are talking about 'now' at the present time. See M.A.K. Halliday (1967/8) 'Notes on Transitivity in English', Journal of Linguistics, 3. 37-81; 3. 199-244; and 4. 179-215.

There is also an interesting discussion of the relation of Functional Sentence Perspective and Textlinguistics in 'Functional Sentence Perspective and Textlinguistics' by Zdena Palkova and Bohunil Palek in Current Trends in Textlinguistics, (ed.) W. Dressler, 1977, Walter de Gruyter and Co.

3. Briefly, Grice is concerned with the distinction between what we say and what is implied by what we say. The latter has to do with Conversational Implicature, which derives from certain maxims which constitute the conditions for successful conversation. These conditions, according to Grice, make for a Co-operative Principle which interlocutors follow. The maxims are given under four headings: quantity, quality, relation and manner. These are given as follows:

Quantity: (i) the speaker's contribution to conversation should be as informative as required but not more than is required.

Quality: the contribution should be true. (i) Speakers should not say anything which they believe to be false; or (ii) for which they lack adequate evidence.

Relation: the speaker's contribution should be relevant.

Manner: (i) the speaker's contribution should avoid obscurity of expression; (ii) ambiguity; (iii) it should be brief; and, finally, (iv) it should be orderly.

For further discussion see H.P. Grice (1975): 'Logic and conversation', in Cole & Morgan (1975: 41-58).

Notes to chapter five

1. An alternative analysis of difficulty in text to the one advocated in this thesis is provided by Roe (1977). Roe's approach is more properly described as semantic than grammatical. The lexis and subject-matter of science texts, he maintains, are the principal factors which cause difficulty. Roe analyses the (abstract) subject-matter of each discipline in to what he calls its 'proto-type' system (P-system). Each system comprises a set of sub-systems between which obtain a network of relations. Corresponding to these underlying P-systems are the various ways we have to communicate about (or describe) them. The latter are what Roe calls the realisation-systems and consist of the verbal, the mathematical and the schematic. The degree of congruence between the P-system and the various realisation-systems accounts for the difficulty in science texts, according to Roe. For example, imprecision in the realisation system of the P-system corresponding to it will introduce a 'noise' factor, which reduces the amount of information which any realisation system can cover. In Roe's view, language constitutes the verbal realisation system. This leads him to ask two questions when assessing scientific text for difficulty:

- (1) is there any loss of information in the process of translating the P-system into the realisation?
- (2) Are there any features of the realisation which do not have correspondence in the P-system - i.e., is there noise as well as information in the text?

Inevitably, Roe maintains, there must be some degree of information - loss in the translation from the underlying P-system to the linguistic description of it in the verbal realisation system. For example, the amount of text is proportionate to the range of systems covered. If this is not the case, then the treatment of certain material in text must be either skimmed or omitted completely. The greater the omission in a text, the less coherent is the text. Viewed in this way, noise is the most contributory factor to difficulty in science text.

Even if we accept Roe's thesis, however, it seems to amount to little more than the assertion that, in language, we can never say everything about the subject covered. Moreover, noise in text, as defined by Roe,

need not always be an obstruction to the reader. For instance, for the purpose of exemplification (as in analogy, etc.), we might expect to find in text material which does not directly correspond to anything in the P-system which underlies it. But such material may be necessary to clarify what is being described, even though it is strictly a digression. It need not lead to difficulty in the text at all. Thus, it seems incorrect to say, as Roe does, that "the amount of text is proportionate to the range of systems covered." Such factors as the purpose of the writer in writing the text, the intended audience for the text and the amount of explanatory material described would all affect what Roe calls the 'systemic range' of a text. The latter may lead to necessary, but strictly speaking, redundant, non-essential material in the text.

REFERENCES

- Botha, R.P. The Justification of Linguistic Hypotheses: A Study of Non-demonstrative Inference in Transformational Grammar, Mouton, 1973.
- Bloomfield, L. Language, Revised edition, Allen and Unwin, 1935.
- 'A set of postulates for the science of language', Language, 2, 153-64
- Bram, V.A. 'Sentence construction in science and engineering texts', The Communicator of Scientific and Technical Information, No.34, January, 1978.
- Chomsky, N.A. Syntactic Structures, Mouton, 1957.
- Aspects of the Theory of Syntax, M.I.T. Press, 1965.
- Chomsky, N.A., & Miller, G. 'Finitary models of language users', Handbook of Mathematical Psychology, Vol. 2, R.D. Luce, R. Bush and E. Galanter (eds.), New York: Wiley, 1963.
- Clark, H.H., & Clark, E. Psychology and Language, Harcourt Brace Janovich, 1977
- Clements, P. The effects of staging on recall from prose, Cornell University, Ph.D. thesis, 1976.
- Crothers, E.J., Paragraph Structure Description: 1 Text, Institute of the Study of Intellectual Behaviour, 1975.
- Dijk T. van. Text and Context: explorations in the Semantics and Pragmatics of Discourse, Longman, 1977 a.
- 'Connectives in Text Grammar and Text Logic, Grammar and Descriptions, Studies in Text Theory and Text Analysis', Teun A. van Dijk and Janos S. Petofi (eds.) Walter de Gruyter. Berlin, New York, 1977 b.
- 'Models for Text Grammars,' 1973, Linguistics 105: 35-68.
- Dixon, R.M.W. What is Language? Longmans 1964.
- Falk, J.S. Linguistics and Language: A survey of Basic Concepts and Implications, John Wiley and Sons, 1978.

- Flores d'Arcais, G.B. 'The Perception of Complex Sentences', Studies in The Perception of Language, W.J. Levelt (ed.), Columbia: Ohio, 1972.
- Fodor, J.A. 'Recognition of Syntactic Structures', The Perception of Language, D.L. Horton and J.J. Jenkins (eds.), Columbia: Ohio, 1972.
- Fodor, J.A., & Bever, T.G. 'The Psychological Reality of Linguistic Segments', Journal of Verbal Learning and Verbal Behaviour, 2965, 4, 414-420.
- Freedle, R.O. 'The Validity of Clausal Processing at the Discourse Level', Discourse Processes, Vol. 1, No. 2, April-June, 1978.
- Garrett, M., & Beyer, T.G., & Fodor, J.A. 'The active use of grammar in speech productions', Perception and Psycholinguistics, 1966, Vol. 30-32.
- Garrett, M. Syntactic Structure and Judgments of Auditory Events: A study of the perception of extraneous noise in sentences, unpublished Ph.D. dissertation, University of Illinois, 1965.
- Glazerfield, Ernst Von 'The Problem of Syntactic Complexity in Reading and Readability', Journal of Reading Behaviour, 1970-1, Nol. 3, No. 2, Spring.
- Goodman, Y., & Greene, J. 'Grammar and Reading in the Classroom', Linguistic Theory: What can it say about reading?, Roger Shuy (ed.), International Reading Association, 1977.
- Grimes, J.E. The Thread of Discourse, Mouton, 1975.
- Grimes, J.E. 'Topic Levels! Theoretical Issues in Natural Language Processing 11, University of Illinois at Urbana-Champaign, 1978.
- Halliday, M.A.K. 'Language structure and language function', New Horizons in Linguistics, London, 1970, Penguin.
- 'Linguistic Function and Literary Style: an inquiry into the language of William Golding's "The Inheritors," reprinted in Literary Style, A Symposium, London and New York: Oxford University Press, 1971, S.B. Chatman, (ed.).
- Halliday, M.A.K., & Hasan, R. Cohesion in English, London: Longman, 1976.
- Hill, A. A. Introduction to Linguistic Structures, Harcourt Brace and Company, New York 1957
- Hirsch, D.H. 'Linguistic Structure and Literary Meaning', Journal of Literary Semantics, 1972, Vol. 1, pp. 80-88.

- Hockett, C.F. A Course in Modern Linguistics, New York, 1958.
- Huggins, A.W.F. Syntactic Aspects of Reading Comprehension, Technical Report, NO. 33, 1977.
- Itkonen, E. 'T.G. and the Philosophy of Science', Transformational-Generative Paradigm and Modern Linguistic Theory, E.F. Koerner (ed.), Amsterdam studies in the theory and history of linguistic sciences, Series 4, current issues in Linguistic Theory, Vol. 1, 1975.
- Jacobson, B. Transformational-Generative Grammar, North-Holland, 1977.
- Johnson-Laird, P.N. 'The Perception and Memory of Sentences' New Horizons in Linguistics, John Lyons (ed.), Pelican, 1970.
- Keen, J. Teaching English: a linguistic Approach, Methuen, 1978.
- Klare, G. The Measurement of Readability, Iowa State University Press, 1963.
- Labov, W. The Social Stratification of English in New York City, Washington, D.C: Centre for Applied Linguistics, 1966.
- 'What is a Linguistic Fact?', in Linguistic Theory 1, The Peter de Ridder Press, 1975.
- Larkin, D. 'Grammar: An Important Component of Reading!', Linguistic Theory: What can it say about reading?, Roger Shuy (ed.), International Reading Association, 1977.
- Leech, G. A Linguistic Guide to English Poetry, Longman, 1969.
- Semantics, Penguin, 1974.
- Leech, G., & Svartick, J. A Communicative Grammar of English, Longman, 1975
- Limber, J. 'The genesis of complex sentences', Cognitive Development and the Acquisition of Language. T.E. Moore (ed.), New York, Academic Press, 1973.
- Lyons, J. Chomsky, Fontana Modern Masters, 1970.
- Semantics, Vol. 2 Cambridge, 1977.

- McMahon, L. Grammatical Analysis as Part of Understanding a Sentence, unpublished doctoral dissertation, Harvard University, 1963.
- Malinowski, B. 'The Problem of Meaning in Primitive Languages', Supplement 1 to Ogden and Richards, The Meaning of Meaning, London: Routledge, 1946.
- Martin, E., Roberts, K.H., & Collins, A.M. 'Short Term Memory for Sentences', Journal of Verbal Behaviour, Vol. 7, 1966.
- Meyer, B.J.F. The organisation of prose and its effect on memory, Amsterdam, North-Holland, 1975.
- McKean, K.O., & Miller, G. 'A Chronometric Study of some Relations between Sentences', Quarterly Journal of Experimental Psychology, 16, 291-300, 1964.
- Morgan, B. 'Further Thoughts about Sentence Construction', The Communicator of Scientific and Technical Information, No. 36, July, 1978.
- Partea, B.H. 'On the Requirement that Transformations Preserve Meaning', Studies in Linguistic Semantics, C.J. Fillmore and D.T. Langendoen (eds.), Holt-Rhinehart-Winston, 1971.
- Pierce, R. Literary Texts: The application of linguistic theory to literary discourse, English Language Research, Birmingham University, 1977.
- Quirk, R., Leech, G., Svartik, J., & Greenbaum S. A Grammar of Contemporary English, Longman,, 1972.
- Sampson, G. The Form of Language, Weidenfeld and Nicolson, 1975.
- Saussure, F. Course in General Linguistics, Jonathan Culler, (ed.), 1974 edition, William Collins.
- Schank, R.C. 'Conceptual Dependency: A Theory of Natural Language Understanding', Cognitive Psychology, 1972.
- Trudgill, P. The Social Differentiation of English in Norwich, Cambridge University Press, 1973.
- Weksel, W., & Bever, T.G. Harvard Centre for Cognitive Studies Progress Report, Harvard University, Cambridge, Mass, 1966.

- Werlich, E. A Text Grammar of English, Quelle and Meyer, 1976.
- Widdowson, H. Stylistics and the Teaching of Literature, London, 1973.
- Winograd T. Understanding Natural Language, New York, Academic Press, 1972.
- Winter, E.O. Connection in Science Material: A Proposal About the Semantics of Clause Relations, CILT Papers and Reports, 1971.
- Replacement as a Function of Repetition: A study of some of its principal features in the clause relations of contemporary English, Ph.D. dissertation for the University of London, 1974. Available in revised form 1977 from University Microfilms International Ann Arbor, U.S.A. Reference 77-70,036.
- A Clause-relational Approach to English Texts: A Study of Some Predictive Lexical Items in Written Discourse, Special Issue, Instructional Science, Vol 6 No 1 January 1977.
- 'Replacement as a Fundamental Function of the Sentence in Context', Forthcoming Forum Linguisticum 1980 Vol 111, either No 2 or 3.
- Towards a Contextual Grammar of English, forthcoming book for Allen and Unwin 1981
- Wittgenstein, L. Philosophical Investigations, Oxford: Blackwell, 1953.
- Yngve, V. 'A Model and an Hypothesis for Language Structure', Proceedings of the American Philosophical Society, 104 (5), 1960.
- 'The Depth Hypothesis', Structure of Language and its Mathematical Aspects, Roman Jakobson (ed.), American Mathematical Society, 1961.

PRINCIPAL REFERENCES

- Burt, M.K. From Deep to Surface Structure,
Harper and Row, 1971.
- Hornby, A.S. The Teaching of Structural Words and
Sentence Patterns, Stages 1 and 2,
Oxford University Press, 1978 Edition.
- Kimball, J.P. The Formal Theory of Grammar,
Prentice Hall Inc., 1978.
- Leech, G., & Svartik, J. A Communicative Grammar of English,
Longman, 1975.
- Winter, E. A Clause-relational Approach to English
Texts: A Study of Some Predictive Lexical
Items in Written Discourse, Instructional
Science, 1977.
- Towards a Contextual Grammar of English, forth-
coming book for Allen and Unwin 1981 especially
the sections on Evaluation, Apposition and
Interpolation in the English Clause.
- Quirk, R., Leech, G., Svartik, J., & Greenbaum, S. A Grammar of Contemporary English, Longman
1972.

Glossary of some of the terms used

Anaphora: A relation of backward reference between a linguistic expression and some antecedent expression in the text. This relation is most commonly illustrated in the use of pronouns:

John bought a new car. It cost him a great deal of money.

Anticipation: A relation of forward reference between an item of the three vocabularies and the following linguistic context. This is part of the contextual function of lexical realisation. (See cataphora).

Apposition: A semantic relation between 'like' grammatical structures - e.g., between adjectives, or between adverbs, or between complements, etc. The semantic relation involves either further specifying (narrowing-down) or amplifying (rephrasing) the meaning of the first appositive. The appositional meaning of the relation is frequently made explicit by the use of adverbials - e.g., namely, more specifically, to be more precise, etc. Apposition may therefore be an implicit or explicit relation.

Cataphora: A relation of forward reference between a linguistic expression and some following expression in the text. A cataphoric pronoun is illustrated in the following matrix-clause:

S1 V C S2
It is interesting that John has decided to resign his post.

Here, there is a cataphoric relation between the pronoun It and the extraposed subject ('John has decided to resign his post').

Clause relation: "A clause relation is the way in which the information of one clause is understood in the light of the other clauses" (Winter, 1974, p.42). Specifically, it is a relation between members, where each member may comprise one or more clauses. The relation may be explicitly signalled by an item from one of the three vocabularies or it may be implicit.

Comparative affirmation: A clause relation where the second member affirms the truth of the first. What is true of X is also true of Y (Winter, 1977).

Competence: Chomsky (1965, p.4) defines competence as "the speaker hearer's knowledge of his language". It refers to the system of rules which the speaker/hearer possesses which define his language. Versions of transformational generative grammar represent attempts to explicate this underlying system of rules.

Depth: This is concerned with the structural expansion of some constituent in the clause. The system of branching which occurs is defined positionally as one of three possible types: left-hand branching; middle branching; or right-hand branching. A clue to the type of branching involved is given by its relation to the main verb.

Evaluation: A crucial relation which subsumes comparative affirmation and comparative denial. It may be thought of as supplying the answer to the general question: "What do you think of X?" More specific forms of this question ask for opinions, evaluations, judgments, attitudes, feelings etc, all of which may strongly predict Reasons or Bases as next sentence.

Grammar: A (formal) grammar, in Chomsky's sense, is some finite specification of the sentences of some particular language. The rules of the grammar must be able to specify either a finite or infinite range of sentences (or strings). In this way, the grammar is said to define the language.

Interpolation: An extension of the adjunct function in the clause, where the structure of the clause is interrupted by an evaluation of some kind.

Language: A (formal) language, within Chomskyan linguistic theory, is described in the following way. "Given some finite set of symbols, V (the vocabulary), a language L over V is a set of finite strings of symbols drawn from V. Although V is finite by definition, L may be finite or infinite." (Kimball, 1973, p.1).

Lexical realisation: This refers to the member of a clause relation which is anticipated or retrospectively realised. Lexical realisation is thus a function of both anaphoric and cataphoric reference in clause relations.

Long subject: The expansion of the normal subject (e.g., small nominal group: 'The tall man') by the addition of postmodifying clauses (e.g., 'The tall man who we saw yesterday when we visited the supermarket').

Mentalism: The philosophical doctrine that human behaviour is in large part undetermined by external stimuli or internal physiological states. This is associated in modern linguistics with Noam Chomsky.

Nominal group: This consists of a noun head which may be premodified and/or postmodified. The nominal group may constitute the structure of any of the following major constituents: subject, object, complement or adjunct.

Performance: The application of the speaker/hearer's competence in either speaking or listening.

Physicalism: The philosophical doctrine "according to which all statements made about a person's thoughts, emotions and sensations can be reformulated as statements about his bodily condition and observable behaviour and can thus be brought within the scope of 'physical' laws ..." (Lyons, 1970, p.97).

Phrase structure rule: These rules are of the form $X \rightarrow Y$ (i.e., rewrite rules) which, when interpreted, means "replace the 'single' element that occurs on the left with the string of symbols that occurs on the right." The output of phrase structure rules provides the input to transformations.

Replacement: This takes the form of direct repetition, deletion, substitution, or a combination of these forms. It is a requirement for acceptable replacement that the repeated structure is lexically distinguished from the structure it is repeating.

Retrospective realisation: The scope of backward reference of lexical realisation. The lexical reference here is thus opposite in direction to anticipation.

Rheme: The 'new' information which is predicated of the subject. In broad terms, this corresponds to the predicate of a sentence, illustrated in the following sentence:

	S	RHEME
	My son-in-law	<u>arrived late for his wedding.</u>

'New' information is the information we offer in reply to a WH-question.

Theme: The 'given' information of a sentence, usually the subject, as in the following sentence:

'Given' information is the established part of the WH-question which demands further information. Thus, the V C element of the above example could have been elicited by a question like this: "What state is the old factory down the road in?" We know about the old factory down the road but we don't know what state it is in.

S. D. NP₁ - Aux - V - NP₂ $\xrightarrow[\text{opt.}]{\text{opt.}}$

(Chomsky, 1957).